М. ТАУБЕ

ВЫЧИСЛИТЕЛЬНЫЕ МАШИНЫ И ЗДРАВЫЙ СМЫСЛ

Computers and Common Sense

THE MYTH OF THINKING MACHINES

by Mortimer Taube

Columbia University Press
New York and London
1961

М. ТАУБЕ

Вычислительные машины и здравый смысл

Миф о думающих машинах

Перевод с английского
С. С. Вениаминова

Предисловие
академика А. И. Берга

Издательство
«ПРОГРЕСС»
Москва 1964

Спецредакция И. Б. Гутчина

Редакция литературы по философским наукам

Предисловие к русскому изданию

Кибернетика как наука об управлении сложными динамическими системами служит задачам повышения эффективности человеческой деятельности в промышленности, экономике, в самых различных направлениях науки. Построение материально-технической базы коммунизма немыслимо без развития и внедрения кибернетики, что зафиксировано Программой КПСС. Отсюда следует, что среди многочисленных проблем кибернетики прежде всего необходимо обратиться к тем из них, решение которых должно удовлетворять насущные потребности нашего общества, строящего коммунизм.

К задачам подлинно актуальным, решение которых не терпит отлагательств, прежде всего следует отнести оптимизацию управления производством, оптимизацию планирования экономики, раннюю диагностику опасных заболеваний, развитие методов и средств программированного обучения, автоматический машинный перевод, раскрытие механизмов наследственности.

Между тем в зарубежной и в отечественной литературе неоправданно много места уделяется дискуссии по вопросу о возможностях кибернетических систем, в частности о так называемом «машинном мышлении».

В ходе дискуссии часто применяются термины, смысл которых понимается неоднозначно. При этом иногда забывают, что для кибернетики характерен подход к изучению мышления лишь в его информационном аспекте. При таком подходе происходит отвлечение от биологической и тем более социальной сущности мыслительного процес-

са, а последний рассматривается лишь как процесс преобразования информации.

Вместе с тем по мере развития кибернетики становится все более очевидным, что для удовлетворения своих непрерывно растущих потребностей человеку во все возрастающем масштабе приходится поручать кибернетическим системам выполнение функций, которые необоснованно считались навсегда закрепленными только за человеческим мозгом. Научная правомерность такого подхода к моделированию и воспроизведению отдельных функций мозга не подлежит сомнению. Однако границы возможного в этом воспроизведении являются спорными и составляют предмет широкой и весьма острой дискуссии. В ходе последней появилось много наносного, претенциозно сенсационного, иногда полумистического. Некоторые горячие головы начинают обсуждать в качестве наиболее актуальных «проблемы» «послушания» машины, «этики» их поведения и т. д. При этом по отношению к изделиям человека — машинам совершенно неуместно применяются термины («мышление» и т. п.), ранее применимые только к самому человеку, что создает путаницу и сенсационную шумиху вокруг машинного воспроизведения элементарных логических операций. Попытки объявить мышлением воспроизведение отдельных формальнологических функций мозга машинами абсолютно неправомерно, так как мыслит не изолированный мозг, а человек, живущий в человеческом обществе, человек с его потребностями, целями и социальными условиями.

Известно, что сознание человека возникло в результате перехода живых существ от животного, стадного существования к коллективной, общественной трудовой деятельности, оно возникло и развивалось как духовное орудие жизни человека в условиях человеческого общества. Таких условий нет и не может быть у машин, а искусственное воспроизведение таких условий нецелесообразно и в обозримом будущем — невозможно.

Сосредоточение преимущественного внимания на вопросе о «думающих» машинах обедняет рассмотрение многих научно-технических и философских проблем, поставленных развитием кибернетики.

В настоящей книге нашли отражение различные точки зрения видных американских специалистов по указанной проблеме.

Работа «Вычислительные машины и здравый смысл. Миф о думающих машинах» принадлежит перу профессора Колумбийского университета Мортимера Таубе — специалиста по программированию и использованию электронных машин в научно-информационном деле. Книга носит остро полемический характер и содержит довольно много критических замечаний, не всегда, впрочем, справедливых.

Большой заслугой автора является детальное исследование значения таких терминов, которые, как он полагает, используются не в однозначном смысле: формализация, возможность, перевод, обучение, смысл и др. В этих терминах анализируются конкретные достижения кибернетики и дается подчас слишком заниженная оценка перспектив ее развития. Автор часто приходит к неоправданно пессимистическим, иногда даже нигилистическим выводам, касающихся возможностей таких практически важных направлений кибернетики, как машинный перевод, обучаемые системы, системы управления в военном деле и т. п. Например, главу о машинном переводе автор резюмирует словами: «...изыскания в области машинного перевода носят характер не истинно научных исследований, а романтического поиска Грааля» (стр. 48).

Специалистам, занимающимся вопросами искусственного разума, Таубе в недопустимом для научной полемики тоне заявляет, что им «... следует либо прекратить болтовню об этом, либо принять на себя серьезное обвинение в том, что они сочиняют научную фантастику с целью пощекотать читателям нервы и в погоне за легкими деньгами и дешевой популярностью» (стр. 58).

Вопреки пессимистическим воззрениям автора несомненно то, что все больше и больше функций головного мозга удается воспроизвести при помощи кибернетических систем. Причем воспроизводятся они со все возрастающей эффективностью. Из подвергшихся «нигилистической» критике направлений работ, связанных с воспроизведением отдельных функций мозга, наибольшее практическое значение имеют работы по обучающимся системам, структурной лингвистике и знаковым системам, нейронным сетям.

Создавая кибернетические системы, человек учится у природы так же, как он учился у нее, создавая орудия физического труда. Однако мозг несоизмеримо сложнее рук и ног и его моделирование еще делает лишь свои пер-

вые шаги. Одной из важнейших способностей мозга, воспроизведение которой имеет большое практическое значение, является способность к обучению. Для создания обучающихся и самообучающихся кибернетических систем большое значение приобретает использование принципов выработки условных рефлексов и других методов изучения головного мозга, разработанных И. П. Павловым. Эти методы помогают в изучении принципов оптимальной организации поисковых действий в неизвестной среде и выявлении принципов оптимального управления сложными системами. Несомненно, машину можно обучить выполнению не чего угодно, а лишь тех операций, на которые она предпрограммирована. Однако это ограничение касается не только машины, но и животных и человека, которые могут обучиться лишь тому, для чего у них есть определенные природные данные. Ценность ставших классическими работ И. П. Павлова по физиологическим основам обучения подвергается М. Таубе сомнению. В то же время превозносятся работы Ч. Шеррингтона, чья чисто субъективистская позиция в вопросах взаимосвязи мозга и мышления подвергалась суровой и справедливой критике со стороны И. П. Павлова.

Роль И. П. Павлова в современной науке трудно переоценить. «Физиология мозга и психология стали подлинными науками только после того, как И. Сеченов и И. Павлов отбросили идеалистические представления о душе, якобы независимой от тела, и вскрыли материальную основу психических явлений»*.

Совершенно неубедителен подход автора и к работам У. Мак-Каллока и В. Питтса. Разработанная ими теория нервных сетей имеет большое значение для исследования принципов управления и переработки информации головным мозгом. Значение этой теории определяется тем, что она служит источником рабочих гипотез, которые проверяются на экспериментальном нейрофизиологическом материале. Логико-математический анализ нервных сетей открывает путь, который в сравнительно недалеком будущем, возможно, приведет к решению такой важнейшей технической проблемы, как обеспечение высокой надежности сложных радиоэлектронных систем, состоящих из

* Л. Ф. Ильичев, «Методологические проблемы естествознания и общественных наук» в сб. «Методологические проблемы науки», изд. «Наука», 1964, стр. 85.

сравнительно малонадежных элементов. Эти работы активно ведутся, в частности, и в Советском Союзе.

В специальной литературе неоднократно отмечались успехи в решении проблемы автоматического перевода (см., например, работы О. Кулагиной, Т. Молошной, И. Мельчука и др. в сборниках «Проблемы кибернетики» № 1—10). Структурная лингвистика и знаковые системы вообще являются теоретической базой для разработки прикладных лингвистических проблем современной кибернетики: речевому управлению различными объектами и процессами, автоматизации информационной службы, автоматизации перевода и реферирования научно-технической литературы, построению информационно-логических машин, конструированию стенографов-автоматов, повышению пропускной способности каналов связи и др. В свете сказанного нельзя согласиться с оценкой, данной автором работам Н. Хомского и И. Бар-Хиллела.

Вместе с тем автор прав в том, что конкретные достижения машинного перевода пока еще достаточно скромны и, по заявлениям некоторых специалистов в этой области, трудно ожидать серьезных, имеющих значение для практики результатов ранее чем через 8—10 лет.

Одной из основных ошибок М. Таубе является непонимание социальной природы человеческого мышления. Вряд ли советскому читателю нужно доказывать, что мысль и сознание — результат трудовой общественной деятельности человека, что человек не может жить и мыслить вне общества.

Автор критикует сторонников «думающих» машин, опираясь на невозможность полной формализации, в то время как главным в данном вопросе является учет социальных аспектов мышления, весьма отличных от тех условий, которые требуют для своего функционирования машины. Впрочем, эта ошибка присуща не только М. Таубе, но и всем остальным американским авторам работ, помещенных в настоящей книге.

Несмотря на отмеченные недостатки, работа М. Таубе содержит много интересного и полезного. Полезны попытки автора дать четкие определения различным категориям и понятиям кибернетики, его призыв к осторожности в оценке ее конкретных достижений; интересны и полезны главы, излагающие отношения в системе человек — машина и др. Что касается философских высказы-

ваний М. Таубе и некоторых его «нигилистических» оценок, то читателю следует подойти к ним критически.

В приложениях помещены переводы трех статей, авторы которых по тем или иным вопросам высказывают отличную от М. Таубе точку зрения.

В своей статье А. Сэмюэль выступает против обеих крайних (как он говорит, «ультра») точек зрения на возможности искусственного разума, то есть возможности воспроизведения некоторых функций человеческого мозга. Он дает обзор достижений в области распознавания образов, программирования игры в шашки и др. На примере игры в шашки показаны общие идеи обучения машин. При этом отмечается, что машина научается обыгрывать в шашки своего программиста благодаря быстродействию и безотказной памяти, позволяющим проделать в короткое время детальный, но лишенный творческой мысли анализ.

Сэмюэль отмечает, что совершенно нечего опасаться превышения машиной разума человека, ибо разум машины предписан человеком и для составления такого предписания необходима более высокая степень мышления, чем для его выполнения. Вывод, к которому приходит автор: «Искусственный разум — не миф и не угроза человеку» (стр. 148).

В статье «Кибернетика сегодня и ее будущий вклад в технические науки» недавно переехавший в США известный английский специалист по кибернетике У. Росс Эшби, отмечая, что целью мира техники является управление, в качестве одной из «путеводных» задач кибернетики считает разработку мозгоподобных принципов управления. При этом он утверждает, что «система может быть подобной мозгу, или разумной, только по отношению к определенной задаче» (стр. 153). Автор подробно рассматривает ограничения, которые накладываются на искусственную мозгоподобную систему (в частности, теоремами К. Шеннона).

При разработке таких систем Эшби рекомендует «...отбросить мысли о неуловимых свойствах мозга, о магических силах мышления и прийти к действительному пониманию работы мозга.... надо начинать с определенной частной задачи, а затем приступать к последовательному созданию механизма управления, сосредоточив внимание на максимизации эффективности» (стр. 158—159).

В целом позиция У. Р. Эшби иллюстрирует давно известное положение о том, что крупные ученые капиталистических стран, стоящие на научных позициях, одновременно находятся и на стихийно материалистической почве.

В статье Поля Армера дается исторический анализ дискуссии о возможностях искусственного разума и делается попытка свести трудности этой проблемы к семантическим трудностям. Автор производит подробный анализ мышления как n-мерного континуума, что само по себе представляет определенный интерес.

Анализ работ, помещенных в книге, показывает, что в термин «искусственный разум» разные авторы вкладывают различный смысл. Но, за исключением немногих особо горячих голов, под этим термином разумеются системы, воспроизводящие отдельные черты функционирования мозга, — системы, которые могут и должны стать надежными помощниками человека в его борьбе за повышение эффективности умственного труда.

В целом книга отражает различные мнения ведущих специалистов США по вопросу о перспективах воспроизведения некоторых функций мозга и будет с пользой прочитана всеми, кто интересуется техническими и философскими аспектами этой актуальной проблемы.

А. И. Берг

1. Введение: машинизация и формализация

Первоначально написанную главу «О научных заблуждениях» пришлось дать в качестве приложения. По замыслу она должна быть введением, однако в ходе работы над текстом выяснилось, что введение в итоге оказалось сводкой выводов и заключений. Поместить такой материал в начале книги — это, по существу, последовать известному принципу «сначала приговор, а потом уже вердикт присяжных». Поэтому рукопись пришлось перекроить, а во введении дать только план изложения и краткое содержание книги.

В течение нескольких лет я со своими коллегами был занят проектированием автоматизированного центра обработки научно-технической информации. По мере того как мы работали над решением возникавших одна за другой проблем и развивали методы преобразования исходной информации к виду, удобному для обработки на машине, мы снова и снова упирались в пределы возможностей машины и в неспособность человека полностью формализовать понятие, идею, смысл. И тогда же мы краем уха прослышали, что где-то пишутся книги о новых видах машин, которые, если их создать, сделают всю нашу работу устаревшей. Периодическое устаревание техники характерно для прошлого и будет иметь место в будущем. С другой стороны, ткацкие станки, установленные сейчас на большинстве наших фабрик, очень похожи на станки, которыми пользовались люди сто лет назад, хотя каждому, кто наблюдает интенсивную вибрацию работающей ткацкой машины, всегда кажется, что должен существовать более простой способ изготовления тканей, и многие изобретатели пытались его найти. Итак, представляется разумным сначала подробно изучить книги о новых информационных машинах, прежде чем отказаться от наших пусть даже ограниченных методов.

В этих книгах описываются машины, способные переводить с одного языка на другой, обучаться именно в том смысле, в каком обучается человек, принимать решения,

короче говоря, выполнять любую интеллектуальную работу, которую способен выполнить человек. В различных главах настоящей книги анализируются предлагаемые доказательства существования таких машин или возможности их реализации в будущем.

Иногда анализ этих доказательств весьма доступен и, надеемся, будет легко понятен читателю. Но подчас задача настолько усложняется, что трудно проследить за построением аргументации. Будем надеяться, что таких примеров мало, и все они оправдываются сложностью и запутанностью исследуемых вопросов. В конце концов, если лиса петляет, то собака должна делать то же самое.

Было бы полезным сейчас объяснить некоторые понятия и термины, которые будут использоваться в следующих главах для обозначения противоположных точек зрения или общей природы исследуемых конкретных явлений. Прежде всего существует противоположность между человеком и машиной, которая предстает как противоположность между физиологическим и механическим. Существуют механистические интерпретации физиологии, так же как существуют и идеалистические интерпретации физики. Но с точки зрения специальных наук наблюдаются разнообразные и очевидные различия между наукой физиологией и наукой механикой и в равной степени между физиологией и электротехникой. Одним из следствий существования таких различий является то, что мы верим физиологу, отвечающему на вопрос из области физиологии, и инженеру-электрику, отвечающему на вопрос из электротехники, но мы не принимаем на веру утверждения инженера-электрика, сделанные им в области физиологии, так как он не столь компетентен в этой области знаний.

Что касается математики, то мы будем рассматривать как противоположные направления интуиционистов, таких, как Брауэр, Анри Пуанкаре, Герман Вейль, и формалистов, таких, как Давид Гильберт и его последователи. Мы принимаем точку зрения, согласно которой ни математика, ни логика, ни синтаксис не могут быть полностью формализованы [1]. С другой стороны, необходимо признать огромное практическое значение дальнейшего развития формальных методов в математике, логике и синтаксисе.

Слово «формальный», встречающееся на страницах нашей книги, заслуживает того, чтобы сделать его содер-

жание понятным каждому. Пока это слово остается специальным математическим термином, невозможно понять, почему должны существовать пределы возможностей машины и каковы эти пределы.

В том смысле, в каком его употребляют математики, логики и теоретики вычислительной техники, термин «формальный» почти синонимичен словам «машинный» или «автоматический». Следовательно, пределы формализации, если они существуют, являются одновременно и пределами машинизации и автоматизации. Ввиду того что отмеченная синонимия слов «формальный» и «автоматический» не может быть почерпнута в словаре или в неспециализированной разговорной речи, имеет смысл рассмотреть, как развивалось специальное употребление этих слов и почему оно служит ключом к проблемам, обсуждаемым в следующих главах.

Прежде всего рассмотрим простой бытовой пример. Сравним костюм, пошитый портным вручную, с костюмом, изготовленным специальной швейной машиной. Во втором
4 случае мы процесс пошива называем обычно автоматизированным, а не формальным. И тем не менее при более глубоком исследовании предмета выясняется, что, прежде чем процесс пошива можно будет автоматизировать, необходимо его формализовать. Один портной может научиться шить у другого путем подражания, и его мастерство будет совершенствоваться благодаря опыту, получаемому в процессе изучения как свойств ткани, так и тонкостей оперирования инструментом и пальцами. Это обучение — не формальный процесс. С другой стороны, прежде чем машину построить, ее необходимо спроектировать, то есть описать в формальных терминах. Машина осуществляет жесткое управление иглой: здесь мы не встретим относительно свободного поведения иглы портного под воздействием упругости ткани или усталости рук. Больше того, с целью ускорения работы некоторые швы придется упростить и исключить специальные процедуры завершения операций. Костюм, изготовленный машиной, окажется дешевле пошитого вручную и во многих отношениях не будет уступать последнему. И лишь при необходимости учитывать специфические особенности ткани (например, у домотканых материалов) ручную работу нельзя заменить машинной. Эту противоположность между ручным трудом и работой машины можно часто наблюдать вокруг

нас. Ручную работу предпочитают машинной не только из-за снобизма, но и потому, что она действительно лучше.

Автоматизировать можно не только некоторые виды ручной работы, но и умственной. И в этом случае связь формального с автоматическим становится еще более ясной. Складывая или перемножая числа, мы руководствуемся определенными правилами, а все, что можно сформулировать в виде правил, в принципе можно автоматизировать. Так, нам известны машины, производящие сложение чисел и даже еще более сложные вычислительные операции.

Существуют некоторые математические задачи, которые трудно понять в терминах одних только правил, и человеку приходится прибегать к диаграммам и чертежам (то есть к геометрическим интерпретациям), чтобы передать содержание правил. Области математики, где пользуются диаграммами, называют интуитивными, или содержащими интуитивные элементы в отличие от формальных. При любой попытке автоматизировать математический вывод эти неформальные, то есть интуитивные, элементы устанавливают предел такой автоматизации. Вот здесь-то и всплывает эквивалентность понятий «формальный» и «автоматический».

Допустим, в центр обработки информации обращается предприниматель с просьбой разработать систему, которая могла бы дать ему определенную информацию о его покупателях. Разработчик и конструктор будущей системы, задавая заказчику ряд вопросов, стремятся заставить его сформулировать свои требования к системе в формальных терминах. Они говорят: «Если вы хотите разделить своих покупателей на две категории — проживающих к востоку от Миссисипи и проживающих к западу от Миссисипи,— потребуется всего два правила, или формальных кода. Если же вам нужно классифицировать покупателей по принадлежности их к тому или иному штату, потребуется 50 правил, или кодов». Если заказчик пожелает разбить своих покупателей на классы по кредитному риску, разработчик может обратиться к таким сведениям, как частота платежей, данные фирмы «Дан и Бредстрит»*,

---

* Роберт Дан (Robert Dun) — американский специалист по коммерческому кредиту, основал фирму вместе с Бредстритом (Bradstreet); с 1893 года выпускал коммерческий бюллетень.— *Прим. ред.*

годовой объем закупок и т. д. Если же заказчик скажет, что всего этого недостаточно и ему хотелось бы знать о намерениях и добросовестности своих покупателей, разработчик ответит ему, что этого формализовать нельзя и, следовательно, машину, которая давала бы ответы на эти вопросы, построить невозможно. Поэтому заказчик должен будет положиться на свое интуитивное суждение. В последнем случае разработчик может немного погрешить и предложить заказчику шкалу добросовестности, скажем, от 1 до 10. Покупателю *A* присваивается добросовестность 0,7, покупателю *B* — добросовестность 1,6 и т. д., разработчику не удастся обосновать эту шкалу, пользуясь какой-либо известной информацией о покупателях, например «историей платежей», потому что полученные числа служили бы мерой истории платежей, а не мерой добросовестности. Короче, добросовестность либо должна непосредственно измеряться так, что ее можно выразить некоторой формальной величиной в нашей системе, либо она остается неформальным элементом, ускользающим от формализации.

Даже в математике иногда выгодно заменять формальный подход интуицией, потому что иной раз формальный анализ требует громадного количества шагов, которые во многих случаях без вреда могут быть опущены в доказательствах и выкладках. Хотя Principia Mathematica* явилась сильным толчком в развитии формальной логики, она все же на своем пути к формальным доказательствам тех положений, которые каждый школьник принимает интуитивно без доказательств, содержит много интуитивных скачков. Куайн в своей «Математической логике» охотно приводит «неформальные доказательства существования формального вывода» [2], потому что, по его словам, задачей его является формализация логики, а не формализация рассуждений о формализации логики.

Если бы трудности чисто формального подхода в математике сводились только к громоздкости и утомительности

---

* Фундаментальный трехтомный труд по основаниям математики А. Уайтхеда и Б. Рассела: Alfred Whitehead, Bertrand Russel, Principia Mathematica, Vol. 1, 1910, XV+ +666 p. (2nd ed. 1925), Vol. 2, 1912, XXIV+772 p. (2nd ed. 1927), Vol. 3., 1913, X+491 p. (2nd ed. 1927). Cambridge, England (University Press).— *Прим. ред.*

вычислений, появление электронных вычислительных машин уже привело бы к полной формализации математики. Так и было со многими проблемами, формальный подход к которым ограничивала громоздкость вычислений и решение которых стало доступным с приходом вычислительных машин. Действительно, успех в этом отношении был так поразителен, что многие стали допускать возможность формализации, а следовательно, и автоматизации решения любой математической проблемы. А это, как было указано выше, недостижимо. Формализация математики предполагает выводимость всех математических утверждений из логических. Однако логическая истина уже сама по себе не формальна.

В последующих главах, посвященных машинному переводу, машинному обучению и лингвистическому анализу, решение многих из рассмотренных выше проблем сведется к ответу на вопрос: может ли быть формализован язык, обучение, мыслительный процесс. Во всех случаях вопрос автоматизации изучается как проблема формализации. Переходить к автоматизации можно только в том случае, если построено формальное правило.

В тех случаях, когда формализация возможна, она не всегда выгодна. Картина, музыкальное или литературное произведение, созданные по формальным правилам, страдают определенной ограниченностью. Мы не уверены даже, что школьника можно научить геометрии, не прибегая к построениям.

Наконец, нетрудно понять, что язык как система содержательных символов, устных или письменных, не есть формальная система и не может быть сведен к ней без разрушения его истинной природы. Вот почему исторически грамматика явилась нормативом, а не описательной наукой. Когда язык формализован, он перестает быть языком и становится кодом. Даже те лингвисты, которые рассматривают речь прежде всего как явление физическое, вынуждены пользоваться неформальным понятием подобия для того, чтобы *разные* звуки классифицировать как *одну и ту же* фонему.

Обычно о формализации математики, лингвистического анализа или обучения говорят в смысле возможности. Ввиду того что даже самые убежденные сторонники формализации и автоматизации согласятся с тем, что ни первая, ни вторая еще не достигнуты, эти две категории дол-

жны рассматриваться (если вообще их рассматривать) как *возможности*. Поэтому в следующей главе мы разберем значение этого чрезвычайно трудного понятия.

ЛИТЕРАТУРА

1. W. V. O. Quine, Mathematical Logic, Cambridge, Harvard University Press, 1951, p. 318.
2. Там же, стр. 321.

2. Возможность как ключ к выбору направлений исследований

В этой главе мы коснемся употребления термина «возможность» по отношению к научным взглядам или исследованиям и понятия «возможный успех» как оправдания программы исследований.

Для пояснения особенностей такого употребления ниже приводится ряд вопросов, содержащих термин «возможность».

(1) Возможен ли машинный перевод с одного языка на другой?

(2) Можно ли создать вечный двигатель?

(3) Можно ли точно измерить положение и скорость электрона?

(4) Можно ли увидеть бога?

(5) Возможно ли внечувственное восприятие?*

(6) Можно ли углубить наши знания о природе горения?

(7) Можно ли установить платформу в космическом пространстве?

(8) Может ли быть истинным высказывание «*p* и не-*p*»?

(9) Можно ли доказать полноту и непротиворечивость аксиом арифметики?

(10) Может ли машина мыслить?

На некоторые из этих вопросов можно сразу ответить «*да*», на другие— «*нет*», а третьи вызовут двусмысленное «*может быть*». Верно также и то, что на одни вопросы довольно легко ответить, тогда как для ответа на другие потребуется тщательный анализ значения не только слова «возможность», но и других слов, содержащихся в вопросе. Например, «*p* и не-*p*» суть противоречие. Следовательно, на вопрос (8) можно ответить «*нет*» с полной уверенностью в том, что никакой прогресс науки никогда не опровергнет этого ответа. Исследовательская программа поиска положительного ответа на этот вопрос была бы вздором.

---

* Автор имеет в виду телепатию.— *Прим. ред.*

На вопрос (6) следует ответить однозначно *«да»*. Горение представляет собой явление природы, в результате постоянного изучения которого в прошлом веке знания о нем все более и более углублялись. Поэтому вполне разумно допустить возможность еще большего углубления наших знаний о процессе горения по мере продолжения его изучения, и соответствующая программа исследований для своего оправдания не требует никаких других условий, кроме компетентности исследователей.

На вопрос (7) ответить несколько труднее. В свете последних достижений ракетостроения и успехов запуска искусственных спутников земли естественно предположить, что дальнейшее развитие систем двигателей, управления, защиты (по своему общему характеру аналогичное углублению знаний о природе горения) позволит ответить на вопрос (7) положительно. Может быть, нам никогда и не удастся установить платформу в космосе. Но, с другой стороны, если экстраполировать современное состояние науки и техники, эта возможность представляет вполне оправданную цель исследований. Установление такой платформы оказалось бы полезным для науки вообще, для дальней связи и систем обороны.

Ежегодно Бюро патентов получает многочисленные предложения конструкций вечного двигателя, но, несмотря на это, большинство людей, даже имеющих элементарные естественнонаучные знания, на вопрос (2) ответили бы отрицательно. Следует заметить, однако, что если вопрос (8) содержит в себе противоречие, то в вопросе (2) ничего противоречивого нет* и, следовательно, отрицательный ответ на него должен следовать не из самого вопроса, а из неявных предпосылок, из которых невозможность (2) можно вывести. Например, если вы принимаете второй закон термодинамики как истину, то на вопрос (2) следует ответить *«нет»*. И если кто-то решил заняться выяснением возможности построения вечного двигателя,

* Здесь трудно согласиться с автором, ибо вопрос (2) включает специальный термин — «вечный двигатель», который должен быть раскрыт на одном уровне формализации с вопросом (8). Если вместо этого термина подставить его описание, то в последнем будет содержаться форма «*p* и не-*p*», а именно неявное допущение второго начала термодинамики и одновременно его явного нарушения. Таким образом, вопрос (8) сведется к вопросу (2), следовательно, различие в характере между обоими вопросами скорее искусственное.— *Прим. перев.*

то этому человеку нужно посоветовать сначала рассмотреть второй принцип термодинамики. Было бы совершенно неразумно, с одной стороны, принимать за абсолютную истину второе начало термодинамики и, несмотря на это, продолжать исследования по созданию вечного двигателя.

То же самое можно сказать и о вопросе (3), хотя противоречие какому-либо физическому закону здесь не так очевидно, как в вопросе (2). Если принцип неопределенности Гейзенберга истинен, то ответ на вопрос (3)— «*нет*». Но некоторые детерминисты пытались возражать, утверждая, что принцип Гейзенберга устанавливает лишь практические трудности точного измерения положения и скорости электрона и что когда-нибудь, возможно, удастся преодолеть эти трудности, а поэтому на вопрос (3) следует ответить «*да*». Тем не менее большинство представителей ученого мира дали бы отрицательный ответ на него, и вряд ли исследователи, поставившие перед собой задачу точного измерения положения и скорости электрона, встретили бы сейчас достаточную поддержку. Стоит заметить, однако, что расхождение мнений по вопросу (3) можно объяснить различием в интерпретациях слова «возможный». Такого различия не возникает при рассмотрении вопросов (2), (6), (7) и (8) и скорее всего ввиду явной определенности ответов на них.

Вопрос (4) относится к типу вопросов, на которые одна категория людей ответит определенно положительно, а другая — определенно отрицательно. Каждая категория для защиты своей позиции будет выдвигать свои аргументы, основанные на различных интерпретациях самого вопроса. Например, те, кто ответил на вопрос (4) «*нет*», могут аргументировать следующим образом: а) бог суть объект бесплотный, нематериальный; б) бог не имеет цвета; в) бог не отражает света; г) глаз может различать цвет световых лучей, отраженных только от материальных тел, поэтому бога видеть нельзя. Те, кто ответил «*да*», могут аргументировать тем, что якобы бог присутствует во всех его деяниях. Дескать, хотя глаза и участвуют в процессе «ви́дения», они выполняют роль лишь стимулов особого рода и не могут сами определить, что они видят. А видит душа, и для души видеть бога значит только то, что бог *является* душе — поэтому видеть бога можно. Нетрудно заметить, что вопрос (4) теперь существенно расширился: на него нельзя ответить определенно, не исследовав природу

таких понятий, как зрение, цвет, материальность, душа и т. д. Но если даже вопрос (4) расширен таким образом и предполагает предварительные ответы на множество других, более трудных вопросов, по-моему, как сторонники *«нет»*, так и сторонники *«да»*, скорее всего, не затрудняют себя таким предварительным анализом понятий. Действительно, ответ на вопрос (4) определяется взглядами, которые не так-то легко поколебать какими-либо результатами научного исследования. Следовательно, как те, кто ответил *«да»*, так и те, кто ответил *«нет»*, сочли бы неразумным посвящать план научных исследований получению ответа на вопрос (4).

Вопрос (5) похож на предыдущий тем, что он тоже заставит отвечающих разделиться на две группы: *«да»* и *«нет»*. Вместе с тем вопрос (5) отличается от (4), ибо нашлись люди, которые защищали и проводили программы исследований, имевшие целью выяснить, является ли ответ на вопрос (5) положительным или отрицательным. Но те, кто уверен в отрицательном ответе, считают такие программы пустым расходованием денег и научных сил. Они скажут, что восприятие по определению чувственно и поэтому вопрос (5) сводится к (8) и на него нужно ответить *«нет»* без всяких исследований. Те, кто отвечает *«да»* или верит в возможность подтверждения существования внечувственного восприятия путем проведения соответствующей программы исследований, отвергнут определение восприятия как чувственного и будут приводить примеры восприятия, возникающего без посредства чувственных стимулов. Их исследовательские программы будут иметь целью обнаружить такие примеры, причем весьма явные.

На вопрос (9) математики отвечали однозначно *«да»* вплоть до 1931 года, когда появилась имеющая большое значение работа Курта Гёделя «О формально неразрешимых предложениях Principia Mathematica и родственных систем» («Ueber formal unentscheidbare Saetze der Principia Mathematica und verwandter Systeme»). В этой работе было доказано, что ответ на вопрос (9) отрицателен. Этот ответ относится не только к формальной логике и математике, но также к вопросу о возможностях машин. К примеру, Тьюринг, отвечающий на вопрос (10) положительно, тем не менее признает, что самые серьезные ограничения возможностей вычислительных машин следует из доказа-

тельства Гёделя*. «Математическая логика располагает множеством результатов, свидетельствующих о наличии ограничений возможностей дискретных машин. Лучшим из таких результатов является теорема Гёделя...» [1]. К аналогичному заключению пришли Нагель и Ньюмен: «Выводы Гёделя имеют отношение также к вопросу, можно ли создать вычислительные машины, способные заменить голову математика. В современном представлении действие такой машины подчиняется фиксированному набору правил, воплощенных в ее конструкции, и используется пооперационный принцип работы. Однако в свете теоремы Гёделя о неполноте элементарной теории чисел существует бесконечное множество задач, которые принципиально неразрешимы этими машинами, как бы сложна ни была их конструкция и как бы быстро они ни работали. Очень может быть, что человеческий мозг — это тоже «машина» с присущими ей ограничениями и с неразрешимыми для нее математическими проблемами. Даже если это так, то человеческий мозг воплощает в себе систему операционных правил, значительно более могущественную, чем у мыслимых в настоящее время машин. Так что в ближайшем будущем не видно перспектив замены человеческого разума роботами» [2].

Итак, будем считать установленной связь между вопросами (9) и (10) в том смысле, что отрицательный ответ на вопрос (9) влечет отрицательный же ответ на вопрос (10). Тем не менее следует признать, что эту связь обычно не замечают. Даже Тьюринг, который, как подчеркивалось выше, безусловно, признает ее, заключает, что на вопрос (10) следует ответить *«да»*. Но все-таки Тьюринг чувствует бессмысленность вопроса (10) в той форме, в какой он задан, и не против заменить эту формулировку, например, следующей: «Может ли машина делать то, что делает человек, когда говорят, что он думает?» В своей статье «Может ли машина мыслить?» Тьюринг вместо этого прямого вопроса рассматривает вопрос о способности машин играть в некоторые игры, требующие, по-видимому, разумного подхода и интеллектуальной работы мозга, если в них играет человек. Выводы Тьюринга очень интересны пото-

* Однако Тьюринг в этой же своей работе замечает, что подобного рода ограничения могут существовать и для человека, то есть нет оснований противопоставлять способности человека способностям машин только на базе теоремы Гёделя.— *Прим. перев.*

му, что он анализирует не только возможность выполнения машиной работы, которую называют «мышлением», если ее выполняет человек, но и значение предположений и догадок в этом направлении для координации программ исследований: «Читателю будет легче ориентироваться в этом вопросе, если я сначала разъясню свои личные убеждения. Рассмотрим сперва более точную форму вопроса. Я уверен, что лет через пятьдесят станет возможным программировать работу машин с емкостью памяти около 10^9 так, чтобы они могли играть в имитацию настолько успешно, что шансы среднего человека установить присутствие машины через пять минут после того, как он начнет задавать вопросы, не поднимались выше 70%. Первоначальный вопрос «могут ли машины мыслить?» я считаю слишком неосмысленным, чтобы он заслуживал рассмотрения. Тем не менее я убежден, что к концу нашего века употребление слов и мнения, разделяемые большинством образованных людей, изменятся настолько, что можно будет говорить о мыслящих машинах, не боясь, что тебя поймут неправильно. Более того, я считаю вредным скрывать такие убеждения. Широко распространенное представление о том, что ученые с неуклонной последовательностью переходят от одного вполне установленного факта к другому, не менее хорошо установленному факту, не давая увлечь себя никакому непроверенному предположению, в корне ошибочно. Не будет ущерба от того, что мы ясно осознаем, что является доказанным фактом, а что предположением. Догадки очень важны, ибо они подсказывают направления, полезные для исследований» [3].

Подчеркивая существенную роль догадок и убеждений в выборе направлений исследований, Тьюринг вскрывает другую двусмысленность, скрытую в понятии «возможность», а именно связь этого понятия с уверенностью в том, что *будущее* ответит положительно на вопрос (10). Можно считать, что некоторое утверждение возможно, не имея каких-либо оснований предполагать, что будущее докажет правильность или ошибочность этого утверждения. Например, можно сказать: «Возможно, что число черепиц крыши дома есть простое число», не имея достаточно серьезных оснований для догадки об истинном числе черепиц. Эта возможность остается таковой до тех пор, пока черепицы не будут пересчитаны, и тогда возможность исчезает, так как число черепиц становится определенным — простым

или нет. Можно сказать, что состояние крыши не меняется от пересчета, и это действительно так. Но убеждение изменяется от состояния возможности до состояния определенности. Это, видимо, касается и убеждения Тьюринга относительно того, что произойдет через 50 лет. Но есть также и существенная разница между употреблением слова «возможность» в только что приведенном примере и функцией убеждения и догадки как руководства при выборе направления исследований. Вряд ли кому-нибудь придет в голову всерьез предложить пересчитать черепицы на крышах домов в качестве программы исследований для выяснения того, какие крыши содержат простое число черепиц. Но тот, кто стоит за исследования, целью которых является создание мыслящих машин, основываясь всего лишь на догадке или убеждении, что их можно построить, должен, конечно, предполагать, что и при отсутствии абсолютной гарантии успеха все же есть существенные основания его предсказывать. Иными словами, простая логическая возможность — это еще не обоснованное убеждение в том, что данное событие должно произойти, и недостаточное основание для создания соответствующего проекта исследований. Существует несчетное множество возможностей, которые никогда не будут реализованы. И, конечно, нельзя оправдывать усилия, вкладываемые в исследования, объект которых принадлежит к числу этих возможностей.

Блестящий отчет о том, к чему приводит чисто математическое понятие возможности, принятое всерьез в качестве руководства к выбору направления будущих исследований, написан несколько лет назад Расселом Мелони [4]. В начале этого отчета говорится о некоем критике, который расправился с автором, сказав: «Конечно, он написал одну хорошую новеллу. Это и не удивительно. В конце концов, мы знаем, что если посадить шесть шимпанзе за пишущие машинки и заставить их нажимать наобум на клавиши, то через миллион лет они написали бы все книги, хранящиеся в Британском музее» [5]. Один из его слушателей, простой человек, глубоко уважающий науку, сопоставляет это заявление с высказыванием математика, который уверяет его, что это «очень удачная популяризация истины, известной каждому школьнику, знакомому с теорией вероятностей» [6]. После таких слов человек решает помочь науке продвинуться вперед. Он

приобретает шесть шимпанзе и шесть пишущих машинок для них, помещает их вместе в одну комнату и ждет результатов. Далее, строго в соответствии с математической возможностью шимпанзе сначала выстукивают «Оливера Твиста», потом прозу Джона Донна, кое-что Анатоля Франса, Конан-Дойля, Галена, избранные пьесы Сомерсета Моэма, Марселя Пруста, воспоминания покойной Марии Румынской и монографию д-ра Уайли о болотных растениях штатов Мэн и Массачусетс» [7].

Математик, которого пригласили взглянуть на все это, несколько смущен. Тот факт, что шимпанзе напечатала нечто содержательное, а не абракадабру, замечает математик, могло быть случайностью, равнозначной выпадению «орла» сто раз подряд при равных вероятностях выпадения «орла» и «решки». Но эксперимент продолжается, шимпанзе никак не хотят печатать бессмыслицы и, кажется, намерены напечатать все книги Британского музея. Математик теряет рассудок, ранит экспериментатора, стреляет в шимпанзе и сам погибает от выстрела экспериментатора. Под занавес последний шимпанзе, истекая кровью, «...падает перед своей машинкой и жалобно глядит на законченную последнюю страницу „Опытов" Монтеня. Нащупав чистый лист бумаги, шимпанзе вставляет его в машинку и одним пальцем выстукивает: „Хижина дяди Тома", Гарриет Бичер-Стоу, гла...—и умирает» [8].

Возможно, это излишняя роскошь — уделять столько места выдумке ньюйоркца, но, как было отмечено Оскаром Уайльдом, природа подражает искусству. Ни один новеллист никогда не выдумал ничего столь фантастического, чтобы его не превзошел какой-нибудь достопочтенный ученый. Сейчас уже стало обычным воздавать должное Жюлю Верну за большинство из наших современных научных достижений. А пример с шимпанзе был недавно использован Россом Эшби в «Анналах математики» Принстонского университета для доказательства разумности машин, причем разумности, которую можно усилить путем постепенного увеличения конечной вероятности осмысленного выходного сигнала при случайном входном:

«Часто отмечают, что любая случайная последовательность, если она имеет достаточную длину; содержит *все* ответы. Ничто не мешает младенцу пролепетать

$$\mathrm{Cos}^2 x + \mathrm{Sin}^2 x = 1$$

или пылинке в луче света протанцевать то же самое высказывание в коде Морзе или каком-либо другом коде. Выразимся более определенно. Если бы каждый из тринадцати символов, составляющих приведенное выше уравнение, являлся одной из букв пятидесятизначного алфавита, то (так как 50^{13} приблизительно равняется 2^{73}) это уравнение можно было бы передать в двоичном коде с помощью 73 двоичных символов. Теперь представим себе сутолоку соударяющихся молекул в одном кубическом сантиметре воздуха. Каждая молекула после соударения с другой отскакивает либо влево, либо вправо, порождая таким образом двоичную последовательность, каждые 73 знака которой в некотором данном коде либо представляют, либо не представляют указанное уравнение. Простые расчеты показывают, что молекулы в каждом кубическом сантиметре воздуха *правильно* выстукивают эту формулу свыше ста раз в секунду. Возражения, что такого, дескать, не может быть, несостоятельны.

Лепет младенца или любой другой случайный процесс способен воспроизвести все что хотите» [9].

Если шимпанзе могут напечатать все книги, хранящиеся в Британском музее, то они в состоянии также перевести их и написать «Хижину дяди Тома» по-арабски, а все русские научные книги по-английски. Некоторые считают это предположение более чем фантастическим, но Уоррен Уивер развивает его совершенно серьезно в своей статье «Перевод», которая положила начало большинству современных работ по машинному переводу.

«Более общие основания для надежд создать машину, которая справилась бы с существенной частью проблемы перевода, следует искать в теореме, доказанной в 1943 году Мак-Каллоком и Питтсом. В ней доказывается, что робот (или вычислительная машина), построенный на регенеративных контурах определенного формального характера, способен вывести любое правильное заключение из конечного числа посылок... а поскольку письменный язык носит логический характер, эта теорема убеждает нас в том, что упомянутая проблема по крайней мере формально разрешима» [10].

Допустим, что теорема Мак-Каллока — Питтса, так же как писательские способности шимпанзе, логически обоснованы. Джон фон Нейман подчеркнул, что такой вывод может состоять из «тысяч или миллионов или совер-

шенно фантастического числа томов» и окажется значительно сложнее системы связи в модели мозга Мак-Каллока — Питтса. Фон Нейман говорит, что проследить все связи в физическом мозге было бы относительно простым способом объяснения формальной сети Мак-Каллока — Питтса и что такая сеть не поможет объяснить действия человека. В точных логических терминах фон Нейман заключает, что результаты Мак-Каллока — Питтса не «эффективны».

Эту мысль можно развить дальше, так как она имеет прямое отношение к ответу на вопрос (1), который служит предметом следующей главы.

Согласно Уиверу, вычислительная машина определенного формального класса способна вывести любое правильное заключение из конечного числа посылок. Если «правильное» означает лишь правильное с точки зрения посылок, то это утверждение, конечно, верно. Если же под «правильным» подразумевается «любое истинное» утверждение, то из теоремы Гёделя следует, что для любой данной вычислительной машины и любого конечного числа посылок существуют не выводимые из них правильные утверждения. Иными словами, на языке теоремы Гёделя для сети Мак-Каллока — Питтса с любым конечным числом посылок и конечным числом правил вывода найдется бесконечный класс истинных высказываний, не выводимых с помощью такой сети. Дело в том, что при повторном изучении статьи Мак-Каллока и Питтса обнаруживаются не только существенные неточности их физиологической модели, но и неполнота логической структуры, обилие неопределенных терминов и даже явные ошибки*.

Работа Мак-Каллока — Питтса [11] имела большое значение, поскольку в ней проведена плодотворная аналогия между логикой функций истинности и моделями нейрологического поведения. Но когда она используется как платформа для защиты утверждения, что действие челове-

* Теорема Мак-Каллока — Питтса утверждает, что любая функция естественной нервной системы, которая может быть логически описана с помощью конечного числа слов, может быть реализована формальной нервной сетью, а следовательно, и воспроизведена машиной. Подчеркнем при этом, что речь идет о потенциальной реализуемости (в смысле потенциальной осуществимости по А. А. Маркову). До настоящего времени какие-либо конкретные доказательства ошибочности этой теоремы в литературе не появлялись.— *Прим. ред.*

ка, включая мышление и естественный язык, можно полностью формализовать, то есть представить в виде набора теорем, выведенных из аксиом Мак-Каллока — Питтса, ее научность становится сомнительной. Больше того, если Мак-Каллок предлагает ее как решение проблемы индукции [12], а Уивер видит в ней официальную гарантию того, что естественные языки можно машинизировать, то эта статья становится научным заблуждением. В двух следующих главах показано, насколько часто встречается это заблуждение.

ЛИТЕРАТУРА

1. A. M. Turing, Can a Machine Think? «The World of Mathematics», ed. by James R. Newman, New York, Simon and Schuster, Inc., 1956, vol. 4, p. 2109.
2. Ernest Nagel and James R. Newman, Goedel's Proof, «The World of Mathematics», vol. 3, p. 1695.
3. A. M. Turing, Can a Machine Think? p. 2107.
4. Russell Maloney, Inflexible Logic, «The World of Mathematics», vol. 4, p. 2262—2267.
5. Там же, стр. 2263.
6. Там же, стр. 2264.
7. Там же, стр. 2265.
8. Там же, стр. 2267.
9. W. R. Ashby, Design for an Intelligence-Amplifier, «Automata Studies», ed. by C. E. Shannon and J. McCarthy, Princeton, Princeton University Press, 1956, p. 217.
10. Warren Weaver, Translation, «Machine Translation of languages», ed. by William N. Locke and A. Donald Booth, New York, John Wiley and Sons, Inc., 1955, p. 22.
11. W. S. McCulloch and Walter Pitts, A Logical Calculus of the Ideas Immanent in Nervous Activity», «The Bulletin of Mathematical Biophysics», ed. by. N. Rashevsky, Colorado Springs, The Dentan Printing Co., 1943, vol. 5, p. 115—133.
12. Information Theory, Third London Symposium, ed. by Colin Cherry, London, Academic Press Inc., 1955, p. 230.

3. Машинный перевод

В настоящее время проводится очень много работ в области машинного перевода. Эти работы, конечно, не были бы оправданы, если бы ответ на вопрос (1) не был положительным. Тот факт, что исследования продолжаются, сам по себе еще не означает, что ответ на этот вопрос положителен, ибо сейчас продолжаются опыты по обнаружению внечувственного восприятия, хотя многие маститые ученые на вопрос (5) отвечают решительно и определенно «*нет*». Далее, как и в вопросах (4) и (5), можно так истолковать некоторые слова вопроса (1), что на него можно будет ответить либо определенно «*нет*», либо определенно «*да*». Если перевод по определению считать не машинальной деятельностью человеческого разума, ответ будет, конечно, «*нет*», потому что вопрос (1) сведется к вопросу (8). Если же перевод истолковать как распознавание кодовых образов, несомых некоторой физической средой, и манипулирование этой средой на уровне различных кодовых образов, то есть так, как машина сортирует перфокарты, то ответом на вопрос (1) будет определенное «*да*». Есть и другие терминологические трудности, связанные с вопросом (1). Тот, кто уверовал в догму XIX века, рассматривающую человека как машину, мог бы ответить на вопрос (1) положительно. Если человек — это машина и он переводит с одного языка на другой, то, конечно же, и машина может делать языковые переводы. Как Мак-Каллок, так и Эттингер, основываясь на этой догме, на вопрос (1) ответили «*да*». Так, первый заявляет: «С тех пор как природа дала нам действующую модель, нет необходимости задавать вопрос, можно ли построить машину, способную перерабатывать информацию так, как это делает мозг» [1]. А другой добавляет: «Таким образом, если нашей целью является создание более надежных и лучше приспособленных к внешним условиям машин, способных бесперебойно работать в течение многих десятилетий, то живые организмы, так сказать, доказывают теорему существования и гарантируют реальность этой цели» [2].

Итак, понятно, что двусмысленность вопроса порождается неоднозначностью не только слова «перевод», но и слова «машина». Но это еще не все: само понятие языка чрезвычайно двусмысленно. В течение XIX и первой половины XX столетий, когда большинство научных дисциплин стремительно развивалось и обогащалось новыми эмпирическими и научными данными, лингвистика выделилась из классической филологии и стала развиваться как наука речи. Ее основные понятия — фонемы и морфемы — относятся к самой речи и речевым структурам. Многие ученые-лингвисты считают, что письменный язык не служит совершенным представлением реальных лингвистических явлений, именно явлений речи. Для этих лингвистов письменный символ с точки зрения лингвистики имеет то же значение, что и звуковой символ. Больше того, многие лингвисты ограничиваются изучением звуков и звуковых сочетаний как чисто физических явлений, не затрагивая нормативных и нефизических проблем грамматики и содержания. Параллельно с этим историческим развитием лингвистики и мотивирующиеся теми же позитивистскими соображениями развивались области математики и логики, которые сами были связаны с чисто формальными «языками», то есть языками, состоящими исключительно из письменных знаков и правил манипуляции ими. В отношении этих языков вопросы смысла вообще неуместны и могут быть отнесены к сфере метаязыков. Эта идея лежит в основе афоризма Бертрана Рассела: в математике мы никогда не знаем, о чем мы говорим и истинно ли то, что мы говорим.

Появление машинного перевода привело к любопытному результату как в лингвистике, так и в формальных ветвях математики. Машинный перевод прежде всего интересуется *смысловой эквивалентностью письменных языков*. Следовательно, он, казалось бы, не может быть ни разделом символической логики, ни разделом научной лингвистики. И тем не менее машинным переводом сейчас занимаются как логики, так и лингвисты. Невольно приходится согласиться с Джорджем Грэбом в том, что «пусть моду выдумал дурак и пусть для дураков, но правит вкусами она умнейших из умов» [3].

Понятно, что двусмысленность терминов «машина», «язык», «перевод» заключается в том, что, выбирая соответствующим образом значения каждого из них, можно

ответить на вопрос (1) либо однозначно *«да»*, либо однозначно *«нет»*. Если бы эти значения были выбраны такими, чтобы ответ был отрицательным, это означало бы, что на погоню за недостижимыми результатами наше правительство и частные фирмы потратили миллионы долларов и привлекли деятельность большого числа ученых. Но каждый, кто пытается выдвинуть подобный аргумент в пользу ограничения исследовательской деятельности, будет обвинен в софистике — и поделом. И наоборот: каждый, кто попытается оправдывать исследования, выбирая значения терминов по своему усмотрению, будет обвинен в софистике точно таким же образом.

Итак, если человек искренне заинтересован в целесообразности исследований в области машинного перевода, он должен быть выше всей этой игры в определения и изучить истинные мотивы, оправдывающие начало таких исследований, определить показатели успеха в зависимости от исходных требований и представить себе реальные (а не только логические) перспективы окончательного успеха.

Часто любят говорить о том, что многие важные научные открытия были сделаны случайно в ходе исследований, посвященных совершенно другим целям. Этот довод, если он справедлив, построен на предположении существования какого-то мистического провидения и может както оправдывать усилия, затраченные на бесплодные исследования. Но отнюдь не следует пускаться в научные предприятия, которые prima facie* обречены на провал в надежде, что авось подвернется какое-нибудь великое открытие. И когда занялись машинным переводом, обращение к провидению стало превалировать.

Недавно на заседании подкомитета специальных исследований Комитета по науке и астронавтике палаты представителей США, посвященном машинному переводу, выступающие уверяли присутствовавших, что «если даже их работа по автоматическому языковому переводу окончится неудачей, они все же надеются сделать вклад в изучение языка и, возможно, в решение проблем запоминания, хранения и выдачи информации». В недавно вышедшей книге по машинному переводу также содержатся подобные утверждения: «На самом деле для будущего

---

* На первый взгляд (*лат.*).— *Прим. перев.*

значение исследований в области автоматического перевода, видимо, определится больше их вкладом в фундаментальное понимание всех уровней структуры языка, чем в создание машины-переводчика» [4].

Мы постараемся на этих страницах как можно тщательнее разобраться с машинным переводом и будем надеяться, что эта тщательность не покажется читателю скучной. Хорошо, что есть книга «Машинный языковой перевод», изданная в 1955 году [5], в которой дан полный исторический обзор машинного перевода вплоть до дня публикации книги. Кроме того, последовательность работ в этой области можно проследить по докладам Национальному научному совету и в разделе «Научная документация по текущим исследованиям и разработкам», посвященном машинному переводу. Есть еще весьма любопытный отчет Йегошуа Бар-Хиллела о состоянии работ по машинному переводу в США и Великобритании.

Сделаем одно последнее уточнение, прежде чем перейдем к подробностям. Все знают анекдот о фермере, впервые увидевшем на ярмарке верблюда и при этом воскликнувшем: «Да нет же такого животного!» Только одно могло бы сделать все то, что вы сейчас читаете, излишним — это *наличие* истинного продукта машинного перевода. К сожалению, все способности машин, предсказанные в 1946 году, когда машинный перевод только зародился, еще не достигнуты и их следует рассматривать как предсказания, а не как рабочие характеристики. Даже Эттингер счел необходимым закончить свой обзор автоматического языкового перевода следующим замечанием: «Хотя некоторые способы гладкого автоматического русско-английского перевода и предлагались в литературе, тем не менее какие-либо опубликованные экспериментальные результаты отсутствуют, не считая нескольких газетных сообщений, которые не получили точного подтверждения в технической литературе» [6].

Во введении к «Машинному языковому переводу» смысл и цель машинного перевода представлены следующим образом: «В большинстве следующих работ неявно предполагается существование взаимно-однозначного соответствия между языком оригинала и языком перевода. Если это допущение правильно, то можно говорить о чисто машинном процессе в широком смысле слова, который, будучи примененным к входному тексту, выдаст на выходе

его перевод и который в результате повторного применения уже к переводу воспроизведет исходный входной текст» [7].

Далее автор пишет, что такой идеальный результат, по-видимому, «в общем случае невозможен», но параграф заканчивается утверждением того, что машинный перевод будет развиваться от автоматических словарей до полностью автоматической системы, ставящей каждому входному тексту «точный и однозначный выходной текст».

Идея машинного перевода родилась в 1946 году во время беседы Уоррена Уивера с А. Д. Бутом, касавшейся возможных приложений вычислительных машин. Позднее Бут и Д. Бриттен разработали «подробный код» для словарного перевода. В 1948 году Р. Ричепз выдвинул идею сокращения объема автоматического словаря путем запоминания корней слов в одной части словаря, а окончаний — в другой на том основании, что раздельное хранение корней слов и возможных окончаний потребует меньшего объема памяти, чем хранение всех образуемых из них слов. Символически это можно записать так: корни + + окончания < корни × окончания. Ричепз, очевидно, не знал, что время отыскания нужного слова в памяти типа «корни + окончания» много больше, чем в памяти типа «корни × окончания». Однако Бут и Локк подчеркивают, что по мере развития и совершенствования запоминающих устройств все меньше будет отдаваться предпочтения раздельному запоминанию корней и окончаний слов.

В 1949 году появляется статья Уоррена Уивера «Перевод», которая дает толчок началу исследований в Вашингтонском и Калифорнийском университетах и Массачусетском технологическом институте. Бут говорит, что нашлось несколько насмешников, но в общем статья Уивера была встречена с энтузиазмом.

Следующим важным шагом в 1950 году явилось введение Эрвином Рифлером предварительного и окончательного редактирования. Функция первого заключалась в подготовке текста к виду, удобному для его машинной обработки, а функция последнего — в придании переводу, выданному машиной, вида, удобного для чтения человеком. Когда же стало ясно, что стоимость предварительного редактирования может намного превысить стоимость самого перевода, большинство авторов работ в этой области, включая самого Рифлера, высказались против предвари-

тельного редактирования. В 1951 году В. А. Освальд и Стюарт Флетчер опубликовали специальное «Предложение по машинизации немецкого синтаксиса», использованное в машине SWAC Национального бюро стандартов. В том же году Йегошуа Бар-Хиллел в Массачусетском технологическом институте начинает заниматься исключительно машинным переводом.

Здесь можно оборвать наш исторический обзор, заметив только, что в 1959 году Бар-Хиллел доложил Национальному научному совету, что «...нет и в ближайшее время не предвидится никакой чисто автоматической методики, которая позволила бы существующим в настоящее время (не обучающимся) вычислительным машинам разрешить полисемию слова «ключ»* в таких фразах, как «Ключ упал в ведро» и «Ведро упало в ключ», в том же смысле, как это сделал бы человек быстро и безошибочно» [8].

Здесь не сказано, что машинный перевод невозможен. Здесь даже не сказано, что чисто автоматический машинный перевод не найдет своего практического применения завтра. Фраза: «нет и в ближайшее время не предвидится» попросту осуждает как бесполезную всю работу коллег Бар-Хиллела, многие из которых стали заниматься машинным переводом по его примеру.

В 1952 году Рокфеллеровский фонд пожертвовал средства на созыв четырехдневной конференции по машинному переводу в Массачусетском технологическом институте. Хотя конференция и не вынесла никакого официального решения, все в общем-то согласились с тем, что работу в области машинного перевода следует продолжать в двух направлениях: 1) конструирование автоматического словаря (время исполнения двенадцать месяцев); 2) немедленный операционный анализ синтаксиса на уровне создания машинной программы. Хотя создать автоматический словарь, возможно, и не так уж трудно, стоит заметить, что ни одного *работающего* словаря за прошедшие после конференции восемь лет так и не было создано. Были написаны программы для словарей специально отобранных терминов и специально отобранных операций, но не существует программы, с помощью которой машина могла

---

* В оригинале «pen» в предложениях «The pen is in the box» (ручка лежит в коробке) и «The box is in the pen» (коробка лежит в детском манеже).— *Прим. перев.*

бы перевести с одного языка на другой любой произвольно выбранный текст.

Работы по машинному переводу продолжались и в 1953 году, но только 1954 год был отмечен двумя выдающимися событиями: 1) демонстрировались способности ИБН-701 переводить с русского на английский, причем использовались словарь на 250 слов и шесть синтаксических правил; 2) вышел в свет первый номер первого тома журнала «Машинный перевод».

В 1955 году вышла книга «Машинный языковой перевод» [5], включившая наиболее важные статьи, написанные в 1955 году, и аннотированную библиографию первых трех номеров «Машинного перевода»; кроме того, в книгу вошли основные работы из этого журнала.

В качестве первой главы книги была перепечатана работа Уоррена Уивера «Перевод», которую, как указывалось выше, можно рассматривать как документ, послуживший толчком к началу исследований в данной области. Эта работа выдвигает три основных положения. Она ссылается на доказательство Мак-Каллока — Питтса как на устанавливающее логическую возможность машинного перевода. Правомочность этого положения уже была нами исследована выше.

Затем эта статья рассматривает отношение между смыслом и контекстом и указывает, что это отношение можно изучать по «статистическим семантическим характеристикам языка». Здесь слово «статистический» понимается не в смысле частоты употребления, а в смысле относительной длины фраз, необходимой для уменьшения неоднозначности. Допущение существенной смысловой неоднозначности контекста порождает новый вопрос — вопрос о минимальном контексте, имеющем вообще любое значение. Бертран Рассел, Уиллард Куайн и большинство современных философов-аналитиков доказали, что именно *предложение*, а не *слово* является наименьшим носителем смысла, содержания. По отношению к машинному переводу это означает, что запоминаемыми элементами должны быть не слова, а предложения. Ингве указывал, что можно построить 10^{50} английских предложений длиной не более чем в 20 слов. Таким образом, если основным носителем содержания служит предложение, то от эффективного машинного перевода нас отделяют 45 порядков. Это, конечно, верно, что части предложений, фраз, слов имеют свой смысл

в обычном значении этого последнего слова, но, с другой стороны, существует много контекстов, в которых смысловое значение какого-либо предложения выводится из содержания целого параграфа. Для всех случаев, когда в качестве единицы носителя содержания мы могли бы взять сочетание слов, меньшее, чем предложение, найдется столько же случаев, когда в качестве такой единицы придется взять группу предложений или параграфов. Те деятели в области машинного перевода, которые решили, что со смыслом можно справиться просто путем машинного запоминания слов, окончаний и синтаксических правил, проглядели существенный элемент семантической стороны современной логики.

В третьем пункте, выдвинутом доктором Уивером, утверждается, что так как во время войны вычислительные машины эффективно использовались в работе с шифровками, они так же эффективно смогут осуществлять автоматический перевод. Это предположение основано на истолковании шифровального дела как разновидности перевода. Ошибочность этой аналогии заключается в том, что в шифровальном деле нет никаких словарей и никаких признаков языка, представленного кодом. Следовательно, одним из первых шагов в шифровальном деле является перестановка кодов самыми различными способами с целью выявления схем и повторений кодовых элементов. Если бы шифровальщик знал структуру языка, на котором составлено перехваченное донесение, и располагал соответствующим толковым словарем, не говоря уже о двуязычном словаре, ему никогда не пришло бы в голову использовать вычислительную машину для выполнения перевода. Что же касается машинного перевода, то здесь предполагают использовать вычислительную машину для перевода с одного известного языка на другой при наличии соответствующего двуязычного словаря. Следовательно, аналогия между шифровальным делом и автоматическим языковым переводом ошибочна по своей сущности. Все три пункта статьи Уивера «Перевод» оказались несостоятельными и не годятся для обоснования доступности машинного перевода.

В книге «Машинный языковой перевод» анализируются такие понятия, как язык, синтаксис, грамматика, семантические единицы, содержание и т. п., и даются описания экспериментов. Эти эксперименты не являются безупреч-

ными, и важность их значения можно признать лишь в том случае, если они опираются на верную теорию языка, содержания и перевода. А что это за верная теория — совсем не ясно из этой книги. Покажем это на нескольких примерах. Риченз и Бут начинают свое описание общих принципов машинного перевода с замечания: «Язык — это последовательность символов, представляющих идеи» [9]. Такое определение мог дать лишь человек, игнорирующий все то, что было сказано по этому поводу за последние 2500 лет. Во всяком случае, по самым современным взглядам символ «Памятник Джорджу Вашингтону» обозначает физическую конструкцию в городе Вашингтоне, а не чью-либо мысль об этой конструкции. Риченз и Бут продолжают: «Самый простой мыслимый письменный язык имел бы один символ на идею» [10].

Это последнее заявление следует рассмотреть в связи со статьей Додда «Стандартный английский язык» [11]. Додд предлагает составить правила перевода для стандартного языка. В этом стандартном языке каждое слово имело бы ровно одно смысловое значение и каждая смысловая единица выражалась бы ровно одним словом, то есть существовало бы взаимно-однозначное соответствие между множеством смысловых значений и множеством слов. Легко показать, что «стандартный английский» Додда (это не что иное, как развитие понятия Риченза — Бута одного символа на идею) приводит к языку, для которого не существует ни синонимии, ни словарей, ни физического мира. Ведь словарь определяет слово либо через синоним, либо через контекст (группу слов), либо с помощью иллюстрации. Если бы каждое слово имело только одно значение и каждое смысловое значение выражалось бы только одним словом, было бы невозможно описать словами, что́ любое данное слово означает.

Взгляды Риченза, Бута и Додда на отношение между символами, с одной стороны, и идеями и содержанием — с другой, отличаются от взглядов, высказанных другими авторами этой книги, не бесплодностью своей, а тем, что их бесплодность выражена более явно. Можно привести еще один пример. Рифлер, рассматривая машинное определение содержания, заявляет, что лингвист, занимающийся машинным переводом, «...не обязан строго придерживаться результатов научного изучения языка. Когда эти результаты отвечают его интересам, он считается с ними.

Но он игнорирует их, если произвольное толкование языковых средств больше его устраивает» [12]. Это поистине reductio ad absurdum* занятия машинным переводом. Это занятие отправляется не от научного изучения языка, а от произвольного определения языка, даваемого с таким расчетом, чтобы машинный перевод выглядел серьезным научным занятием. Заявление Рифлера поддерживает ранние взгляды, заключающиеся в том, что машинный перевод требует такой перестройки логики и лингвистики, чтобы придать своей экстравагантности добропорядочный вид.

С 1955 года обо всех работах по машинному переводу сообщалось в докладах Национального научного совета, а Бар-Хиллел оценивал результаты. Все, что делалось в области машинного перевода, представляет собой множество попыток разрешить смысловую неоднозначность путем формального синтаксического анализа различных естественных языков: английского, русского, немецкого, китайского и др. Формальный анализ нужен был для того, чтобы сделать возможным составление программы для вычислительной машины, которая могла бы эффективно переводить текст, поданный на вход, с одного языка на другой. Обсуждалось также много вопросов, связанных с языком-посредником, то есть общим машинным языком, который мог бы посредничать между всеми естественными языками в процессе машинного перевода. Есть два пути оценки этих работ. Проблему машинного перевода можно рассматривать как формальную проблему и ставить логический вопрос о возможности решения этой проблемы. Или же машинный перевод можно изучать как практическую задачу, и как таковую ее можно кратко сформулировать следующим образом: окажется ли использование автоматических словарей и программирование вычислительных машин для получения приемлемого перевода более дешевым, чем воспитание и обучение переводчиков или даже воспитание целого поколения многоязычных переводчиков?

Многое из того, что было сказано в предыдущей главе относительно формальной возможности предположений, высказанных в вопросах (9) и (10), о теореме Гёделя и рассуждении фон Неймана об эффективности выводов Мак-Каллока — Питтса, применимо здесь и к вопросу о теоретической возможности машинного перевода. Мы

---

* Сведение к абсурду (*лат.*).— *Прим. перев.*

уже это заметили, когда касались заявления Уивера о том, что проблема машинного перевода формально разрешима.

Но даже без ссылки на труднопонятные и заумные аргументы с помощью простого рассуждения можно установить формальную невозможность машинного перевода. Формальная система должна использовать принципы подстановки и эквивалентности. Если язык A формально (автоматически) отобразим в язык B, то на некотором уровне должно существовать взаимно-однозначное синонимическое соответствие между элементами A и элементами B. Эта мысль проскользнула еще у Бута (см. выше). Можно утверждать, что взаимно-однозначное синонимическое соответствие должно существовать между словами, фразами, предложениями, синтаксическими структурами, параграфами языков A и B и т. д. Вопрос состоит в следующем: истинны ли эти утверждения? Никто всерьез не допускает существования такого взаимно-однозначного соответствия ни для какой пары естественных языков или даже для естественного языка и формального, как, например, арифметика или исчисление высказываний.

При отсутствии взаимно-однозначного синонимического соответствия между определимыми элементами одного языка и определимыми элементами другого человек все же делает перевод. Известно, что два перевода одного и того же текста, выполненные двумя переводчиками, могут оба считаться качественными, хотя и будут отличаться друг от друга по своему словарному составу, фразовым образованиям и вообще по общей синтаксической структуре. Иным покажется, что такой недетерминированный способ перевода, используемый человеком, то есть способ, основанный на неэквивалентности языков, доказывает на первый взгляд формальную возможность осуществления машинного перевода без такой эквивалентности элементов одного языка элементам другого. Вся трудность здесь кроется в понятии «формальный». Сейчас нам придется вернуться к теме, затронутой в первой главе, и обратиться к историческим противоречиям в математике между *интуиционистами* и *формалистами*.

В настоящее время благодаря теореме Гёделя и другим аналогичным результатам в логике и математике положение таково, что интуиционисты нападают, а формалисты защищаются, даже несмотря на обилие работ по современной логике, представляющих попытки развить различ-

ные ограниченные формальные системы, допускаемые теоремой Гёделя, и рассмотрения методов, приводящих к ним. Работы З. Харриса и Н. Хомского представляют собой попытку рассматривать естественный язык как формальную систему, хотя Альфред Тарский (формалист) считает это невозможным. Поскольку язык включает всю математику и поскольку есть доказательство того, что вся математика не сводится к формальной системе, следует, что естественный язык несводим к формальной системе. В какой связи все это находится с машинным и немашинным переводом, сейчас будет пояснено. Мы считаем, что машине чужда интуиция. Следовательно, если машине и суждено переводить, то она будет делать это формально. С другой стороны, и самые строгие формалисты никогда всерьез не отрицали интуитивных способностей человека даже тогда, когда они хотели показать, что математический вывод можно проводить без интуиции. Слово интуиция употребляется здесь в смысле ничуть не более таинственном, чем обычные слова — опыт, ощущения. Сказать, что человек переводит неформально,— значит подчеркнуть, что в каждом акте перевода он неформально пользуется своим арсеналом опыта и чувств. Кое-кто мог бы возражать против этого, утверждая, что словесное выражение опыта и чувств происходит через их формализацию, а тогда уже работает аргумент Мак-Каллока и Питтса: все выразимое (в отличие от усваиваемого) можно моделировать сетью Мак-Каллока — Питтса. Окончательный ответ на этот аргумент состоит в следующем: не существует ни малейшего намека на доказательство эквивалентности опыта всему словесно выразимому. На этой эквивалентности может настаивать лишь тот, кто отрицает свою принадлежность к человеческому роду, кто никогда не слушал музыку, не имеет представления о живописи, никогда не влюблялся и не был ничем глубоко захвачен. Итак, можно сделать вывод, что машинный перевод в формальном смысле невозможен, потому что невозможен формальный перевод с одного естественного языка на другой. Таким образом, мы подошли к следующей проблеме: возможен ли машинный перевод как практическое пособие в получении относительно приемлемого выражения в одном языке идей, заданных в другом языке.

Допустим, что способности человека переводить интуитивно соответствует способность машины переводить

апроксимативно. Здесь апроксимация — это строгий математический процесс приближения к пределу, который в принципе может никогда не достигаться. Например, если смысл определяется контекстом, можно построить автоматический словарь, в котором запоминаются не отдельные слова, а сочетания из двух-трех слов, фразы, целые предложения и т. д. Если, как и считает большинство современных логиков, основным носителем содержания служит предложение, а блоки памяти современных вычислительных машин имеют объем на 45 порядков ниже требуемого для составления словаря из предложений и если в общем память вычислительных машин растет со скоростью один порядок в десятилетие, то вычислительную машину с требуемым объемом памяти можно ожидать через 450 лет *.

Если говорят, что некоторые слова и фразы имеют свой смысл независимо от их вхождения в то или иное конкретное предложение, то можно также сказать и о том, что иногда в качестве единиц запоминания приходится брать контексты большие, чем предложение. Очевидно, мы не получим никакой пользы и никаких практических результатов, если будем пытаться решать проблему машинного перевода «в лоб», путем увеличения числа запоминаемых контекстов в словаре.

Другой «практический» подход к проблеме машинного перевода состоит в попытке составить программу вычислительной машины, способную определять синтаксическую структуру предложения путем анализа не контекста, а таких характеристик предложения, как порядок слов в нем, окончания слов, структура фраз и т. д. Стоит заметить, что этот подход содержит порочный круг, так как более разумно предположить, что части речи, порядок слов, окончания и вообще синтаксический характер частей предложения определяются содержанием предложения, а не наоборот — не содержание предложения определяется синтаксическими свойствами его частей. В обычных учебниках грамматики анализ предложения

* С подобным расчетом трудно согласиться, так как он базируется на предположении о неизменном темпе технического прогресса. В действительности темп технического прогресса все больше и больше ускоряется; см. об этом в статье Л. Ф. Ильичева «Общественный прогресс и философия» в сб. «Наука и человечество», т. 2, 1963, стр. 31—56.— *Прим. ред.*

проводится для того, чтобы показать, *как* выражен смысл, а не для установления того, *каково* содержание. Без предварительного понимания содержания предложения такого грамматического анализа проводить нельзя. Например, мы исходим из понимания смысла данного слова, когда говорим, что оно является подлежащим предложения. Мы не смогли бы узнать, какой символ служит подлежащим предложения, если бы не знали значений символов.

На прямой и ясный вопрос о практической (в отличие от теоретической) реализуемости машинных программ синтаксического анализа и определения смысла естественной реакцией было бы поискать такие программы. Поиск не дал положительного результата, но зато можно категорически заявить, что через 12 лет после опубликования работы Уоррена Уивера * никакой практически полезной программы машинного перевода не существует. Как было замечено выше, нет даже приемлемого автоматического словаря. И это совсем не от того, что в этом направлении не работали. Бар-Хиллел подсчитал, что ежегодные затраты на исследования в области машинного перевода, не считая стоимости эксплуатации машин, составляют приблизительно 3 млн. долларов. По сравнению с многомиллиардным федеральным бюджетом эта сумма может показаться незначительной, но ее достаточно для открытия университета средней величины и для содержания 300 переводчиков с окладом 10 000 долларов в год или обучения переводу 300 студентов с такой же стипендией. Через несколько лет мы имели бы большой контингент компетентных переводчиков, особенно если все это совместить с текущими ежегодными затратами на обучение иностранным языкам. Более того, обучение иностранным языкам могло бы снять необходимость в переводе.

Неверно было бы предполагать, что если машины смогут переводить, то они будут делать это бесплатно. Вычислительные машины очень дорого стоят, а программирование и «откладка» программ тем более. Кроме того, кодирование печатного текста — это дорогая операция. Даже самые горячие сторонники машинного перевода заявляли, что машинный перевод не будет экономически доступным до тех пор, пока не будет создано «читающее устройство», способное автоматически преобразовывать печатный текст

* Речь идет о статье Уоррена Уивера «Перевод».— *Прим. перев.*

в последовательность пробивок на перфокартах или в какой-нибудь другой код, легко воспринимающийся машиной. Имеются читающие устройства, которые могут оперировать со шрифтами стандартной формы, стандартного размера и при строго определенном положении букв. Но никто еще не придумал и не создал читающего устройства, могущего правильно воспринимать любой шрифт. Почта вполне может обратиться к своим клиентам с просьбой употреблять на конвертах надписи стандартного типа и положения в обмен на улучшение качества обслуживания. Однако нам, по-видимому, более разумно попросить русских снабжать их публикации краткими тезисами на английском языке вместо того, чтобы просить их стандартизировать шрифты и типографские операции для облегчения задачи наших конструкторов читающих устройств.

Если вопреки сенсационным заголовкам популярной прессы и даже сообщениям серьезных научных журналов машинный перевод в действительности не делает успехов, то позволительно задать вопрос: есть ли достаточные основания предполагать, что в будущем мы дождемся этих успехов? Короче говоря, где доказательство того, что ближайшее будущее принесет нам экономически доступный машинный перевод?

В этом отношении откровением служат выводы и рекомендации Бар-Хиллела, так как они являются результатом опосредствования всех основных работ по машинному переводу, контролируемых правительством; в конце концов Бар-Хиллел был одним из первых энтузиастов и зачинателей работ в этой области. Он говорит: «Пока что исследования в области машинного перевода должны ограничиваться оказанием переводчику помощи со стороны машин, при этом нужно стремиться к увеличению и совершенствованию этой помощи» [13].

Помощь машины проявляется в различных видах, например замена гусиных перьев автоматическими, замена механических пишущих машинок электрическими, пользование словарем вместо того, чтобы все запоминать самому, и даже введение электрического освещения вместо свечного. Эти виды машинной помощи действительно существенно повышают эффективность работы переводчика.

Не нужно думать, что это слишком вульгарное толкование слов Бар-Хиллела и что он, видимо, имел в виду помощь вычислительной машины переводчику. Вряд ли

можно утверждать, что вычислительные машины могут все это делать, и Бар-Хиллел сам не рекомендует обращаться к ним. Он продолжает: «Частично автоматизированный переводный центр выиграл бы в экономическом отношении, если разработать для него надежное читающее устройство и сконструировать специализированную переводящую машину» [14]. Бар-Хиллел не настаивает на использовании любой существующей вычислительной машины, а предлагает «сконструировать специализированную переводящую машину». Теперь позволительно спросить: что же это за специализированная переводящая машина? Сделал ли ее кто-нибудь когда-нибудь? Знает ли Бар-Хиллел, как ее построить? Если же ее можно спроектировать и построить, то будет ли она, по определению, выдавать адекватные переводы? Здесь можно называть сортировальные машины, машины для обработки информации, но попробуйте такую машину заставить не сортировать карты, а делать что-нибудь другое! Можно поезд называть самолетом, но он все же останется поездом и не станет от этого летать. Излишне уже говорить, что существуют надежные читающие устройства, совсем не помогающие переводчикам, и что не существует никакой экономической стороны частично автоматизированного переводного центра. Можно было бы создать то, чего не существует, но трудно понять, каким же образом может выиграть то, что не существует.

Как уже отмечалось выше и будет снова и снова подчеркиваться в этой книге, по мере того как мы будем развенчивать некоторые модные научные заблуждения, каждое такое заблуждение нужно анализировать, пока в этом есть необходимость, то есть его нужно похоронить в литературе, чтобы потом на это ссылаться.

У Бар-Хиллела в «Отчете о состоянии работ по машинному переводу в Соединенных Штатах и Великобритании» нельзя найти доказательства существования практического машинного перевода, хотя автор считает, что исследованиям в этой области следует придавать более практический характер, повернуть их от теории в сторону практических целей. По его словам, цели, к которым нужно стремиться, не сделают эту область интересной, и ни он, ни кто-либо другой, кто ищет «фундаментальных» проблем, не найдут их здесь. Бар-Хиллел знает, что реальность достижения практической цели, то есть действительного создания экономически доступного семантически тождествен-

ного машинного перевода, могут оценить многие из тех, кто не последовал за ним в сокровенные уголки логики, и вполне понятно, что ему не хочется подвергать такой оценке свою собственную работу. В 1955 году Бут привел несколько цифр, дающих представление о стоимости пословного перевода с использованием аппаратуры, оперирующей перфокартами. Эти цифры превысили вдвое стоимость перевода, выполненного человеком. С того времени мало кто из исследователей удосуживался изучить эти стороны практической проблемы машинного перевода.

Недавно один из старейших и крупнейших центров деятельности в области машинного перевода опубликовал обзор, охватывающий период исследований с мая 1956 по июнь 1958 года [15]. Один из параграфов этой книги называется «Конструкция практического автоматического переводчика с русского на английский» [16], а другой— «Несколько слов об экономическом аспекте машинного перевода» [17]. Первый параграф является типичным примером большинства статей в данной области именно потому, что не дает никакой конструкции, а просто затрагивает некоторые из проблем, встающих на пути к такой конструкции. Вся его практичность состоит в проведении расчета желаемых скоростей оперирования словарем. Но эта статья не дает никаких стоимостных оценок машинного перевода в сравнении с естественным переводом, следовательно, практическое значение ее равно нулю. Во второй работе все же сделана попытка произвести расчет сравнительных стоимостей пословного перевода, выполненного человеком, и машинного перевода. В ней указывается, что полная стоимость машинного перевода складывается в основном из 1) стоимости подготовки текста к вводу его в машину; 2) стоимости поиска слов в словаре и составления первого варианта пословного перевода; 3) стоимости логической обработки этого варианта с целью улучшения его качества и 4) стоимости получения готового оттиска улучшенного перевода. Грубая оценка этих стоимостей показывает, что без специального электронного читающего устройства машинный перевод будет обходиться вдвое дороже естественного перевода и в силу этого «не может быть оправданным с точки зрения стоимости». По-видимому, основная доля стоимости падает на читающее устройство, а не на машину переводчика. Это, конечно, меняет проблему «является ли машинный пере-

вод практически достижимой целью» на проблему «является ли читающее устройство, способное воспринимать шрифт любых размеров и типов, практически достижимой целью». Может быть, и является, хотя нет действительных оснований для положительного ответа. Во всяком случае, практическая неразрешимость машинного перевода допускается, и исследования в области машинного перевода в отличие от исследований по разработке читающего устройства вряд ли перспективны. Здесь проблема читающего устройства служит для отвлечения внимания. Она заслоняет глаза тем, кто ответствен за исследования в области машинного перевода, на тот основной факт, что нигде в литературе нет научного обоснования возможности инженерной реализации машинного перевода. При отсутствии такого обоснования и в свете известной неформальности языка и смысла изыскания в области машинного перевода носят характер не истинно научных исследований, а романтического поиска Грааля.

ЛИТЕРАТУРА

1. W. S. McCulloch, The Design of Machines to Simulate the Behavior of the Human Brain, «IRE National Convention (Symposium), 1955, p. 240.
2. A. G. Oettinger, там же, стр. 242.
3. G. Grabbe, in «The Library».
4. A. G. Oettinger, Automatic Language Translation, Cambridge, Harvard University Press, 1960, p. vii.
5. «Machine Translation of Languages», ed. by William N. Locke and A. Donald Booth, New York, John Wiley and Sons, Inc., 1955.
6. A. G. Oettinger, Automatic Language Translation, p. 346.
7. «Machine Translation of Languages», p. 1.
8. Y. Bar-Hillel, Current Research and Development in Scientific Documentation, Washington, «National Science Foundation», 1959, No. 5, p. 64.
9. «Machine Translation of Languages», p. 24.
10. Там же, стр. 24.
11. Там же, chap. 10, p. 165—173.
12. Там же, стр. 137.
13. Y. Bar-Hillel, Report on the State of Machine Translation in the United States and Great Britain, Jerusalem, Hebrew University, 1959, p. 38.
14. Там же, стр. 39.
15. «Linguistic and Engineering Studies in the Automatic Translation of Scientific Russian into English», Seattle, The University of Washington.
16. Там же.
17. Там же.

4. Обучающиеся машины

Около года назад автору посчастливилось присутствовать на встрече специалистов по обработке информации, состоявшейся в одной организации в Пауфкипси*. Разговор зашел об обучающихся машинах. Большинство присутствующих не сомневались в том, что машины, обладающие способностью обучаться, скоро будут созданы. Когда же был поставлен вопрос о природе обучения у человека и машины и идентичны ли эти свойства у первого и у второй, можно было наблюдать любопытный факт. Все были в основном единодушны во взгляде на природу обучения у машины, однако, что касается природы обучения у человека, то по этому вопросу возникли существенные разногласия. Все были согласны с тем, что термин «обучение» в применении к поведению человека все же остается неясным и неопределенным, несмотря на попытки психологов создать теорию обучения. Из всего сказанного напрашивается любопытный вывод. Раз термин «обучение» не имеет точно определенного смысла, будучи примененным для описания поведения человека, и имеет точный смысл при описании действия машины, то почему бы не взять определение этого термина, принятое для машины, и распространить его на деятельность человека. Другими словами, отождествление человека с машиной достигается не путем моделирования свойств, присущих человеку, а путем приписывания человеку ограничений, свойственных машине. Некоторые считают, что не имеет большого значения, завышаем мы или занижаем уровень сравнения, коль скоро нашей целью является установление тождества между человеком и машиной. Но если, с одной стороны, это отождествление вызывает чувство отвращения, а с другой — совершенно ненаучно, то понимание того, каким образом развивалась эта концепция и почему многие отка-

* Пауфкипси (Poughkeepsie) — небольшой город на восточном берегу Гудзона между Нью-Йорком и Олбани.— *Прим. ред.*

зываются ее оспаривать, поможет нам раскрыть ее характер как характер научного заблуждения.

В отношении машины слово «обучение» употребляется в смысле «улучшения качества ее функционирования на основе прошлой истории ее работы». Для многих существующих машин можно составить программу, которая изменяла бы состав своих команд. Этого можно было бы добиться путем устройства обратной связи выхода программы с какой-либо ее предшествующей частью. Некоторые теоретики пошли дальше этого. Они утверждают, что можно построить машину, которая даже без подробной программы изменяла бы свое поведение в сторону повышения вероятности получения определенного выхода, если критерий желательности этого определенного выхода вложен в машину в ее исходном состоянии.

Машинное обучение, если оно вообще имеет место, необходимо должно быть процессом формальным в математическом смысле этого слова. Это также справедливо, как уже указывалось выше, и в отношении машинного перевода, и здесь, как и раньше, при обсуждении машинного перевода, мы должны задать вопрос, есть ли какие-нибудь основания считать обучение человека просто формальным процессом. Если даже процесс обучения человека частично формален, то он все же есть нечто существенно большее, чем просто формальный процесс, а следовательно, слово «обучение» применяется к человеку и к машине не в одном и том же смысле. Тождество «обучение человека = обучение машины» порочно в том же самом смысле, в котором порочно тождество «все числа = целые числа». Легко показать порочность последнего тождества — ведь очень просто привести примеры чисел, которые не являются целыми. Теперь нужно придумать пример человеческого обучения, не являющегося формальным или машинальным.

Выполнение этой задачи можно начать с замечания, что попытка объяснить выработку привычек и обучение человека с помощью условных рефлексов представляет собой еще один пример широко распространенного научного заблуждения. Сэр Чарлз Шеррингтон* — вот кто может

* Чарлз Скотт Шеррингтон (1859—1952)—крупный английский физиолог, концепции которого о природе человека за их идеалистический характер получили резкое осуждение со стороны И. П. Павлова. В работе «Человек по своей природе» Шеррингтон утвер-

подтвердить, что, хотя поведение живых организмов и можно изменять, подвергая их специально построенной схеме экспериментов, рефлексы в физиологическом смысле этого слова не являются условными, или изменимыми. Короче говоря, выражение «условный рефлекс» противоречиво, поскольку физиолог отличает рефлекторную деятельность от других типов нервной деятельности на основании того, что рефлексы не могут быть условными. И это ничуть не более странно и не более таинственно, чем отрицание ламарковского принципа наследования благоприобретенных признаков. Шеррингтон явно и категорически отделяет рефлекторное поведение от проявления привычек по следующему признаку: привычка приобретается и может изменяться, а рефлекс нет. Если Павлов хотел выработать условный рефлекс, то ему следовало взять что-нибудь попроще, вроде коленного рефлекса, а не слюновыделение у собаки. И. П. Павлов в своих знаменитых опытах не изменил рефлекса выделения слюны. Он попросту изменял его сознательное стимулирование. Если у собаки выделяется слюна, когда она предвкушает пищу, то слюнотечение у нее возникнет и в том случае, если звук колокольчика заставит ее предвкушать пищу. По-видимому, этот опыт доказывает, что рефлекс не меняется, даже если существенно изменить его сознательную стимуляцию*.

Несмотря на то что эти ссылки на работы Шеррингтона и Павлова, казалось бы, совершенно не относятся к нашей теме, они тем не менее сделаны с целью пояснить нашу точку зрения на один факт. Тьюринг в одной из своих работ говорит, что он не намерен уделять много внимания изучению математических аргументов против способности машин мыслить (или обучаться) потому, что тот, кто склонен отрицать эти способности машин, не принадлежит к тому сорту людей, которые пользуются математическими аргументами или которых эти последние могут убедить. Тьюринг предпочитает сосредоточить огонь своей критики на тех, кто возражает против его выводов с религиозных

ждал, что мозг и сознание представляют две друг с другом не связанные системы, что деятельность мозга не имеет отношения к возникновению сознания (см. C h. S h e r r i n g t o n, The Brain and its Mechanisms, Cambridge, University Press, 1934; его же, Man on his Nature, Edinburg, 1955).— *Прим. ред.*

* Автор допускает грубую ошибку, не проводя различия между безусловными и условными рефлексами.— *Прим. ред.*

или моральных позиций. Таким образом, Тьюринг поддерживает ни к чему не обязывающее и ложное допущение, выдвигаемое многими другими энтузиастами вычислительной техники и состоящее в том, что вся научная литература является (по выражению Уильяма Джеймса) консервативной, а сам Тьюринг принимает сторону тех, кто принижает значение жизни и возвеличивает технику.

Математический и физиологический анализы имеют главное значение при рассмотрении машинного обучения. И основная доля той псевдоматематики и псевдофизиологии, которая накопилась вокруг этого вопроса, была порождена теми, кто отождествлял машинное обучение с обучением человека. Например, «Исследования автоматов», помещенные в «Анналах математических исследований», содержат две статьи Альберта Аттли под заголовком «Машины условной вероятности и условные рефлексы»*. Эти статьи вводятся следующим общим заявлением:

«Важной характеристикой поведения живого организма является тот факт, что одну и ту же двигательную реакцию животного можно вызвать множеством различных воздействий, или образов мира, внешнего для этого животного. Для него все эти образы в некотором смысле похожи друг на друга. В настоящей работе предполагается, что это «сходство» образов представляется двумя известными математическими отношениями: первое взято из *теории множеств* — это отношение *включения*; второе — отношение *условной вероятности*. Из этих двух отношений выводятся принципы конструирования машины, реакция которой на стимулы во многих отношениях сходна с реакцией животного ...» [1].

Разбирая сначала отношение включения, Аттли добавляет: «... и в теории множеств не рассматривается никакого другого отношения между элементами множества, кроме отношения включения в него» [2]. В действительности же в теории множеств отношение между множеством и его элементами является не отношением «включения», а отношением «членства». А это совершенно разные вещи**. Отношение «включения» связывает не множество с его

* См. статьи А. Аттли в сб. «Автоматы», Издательство иностранной литературы, 1956, стр. 326—361.— *Прим. перев.*

** Смешение отношений «включения» и «членства» («принадлежности») приводит к возникновению распространенной ошибки, состоящей в отождествлении элемента *a* множества с одноэлементным

элементами, а множества между собой. Заявление Аттли о том, что он выводит принципы конструирования машины из этого положения теории множеств, верно лишь в том смысле, что ложное высказывание влечет любое высказывание. Из такого описания теории множеств он мог бы спокойно вывести, что «луна сделана из зеленого сыра». Путаница множеств и элементов множеств царит у Аттли и при последующем рассмотрении им условной вероятности и условных рефлексов, так что невозможно понять, о чем он говорит. Очевидно, если что-то является правильным из «политических соображений» в отношении условных рефлексов и машинного обучения, то это совсем не обязательно должно быть верным с точки зрения математики, даже если это напечатано в «Анналах математических исследований».

Наряду со слабой математикой Аттли можно привести примеры слабой физиологии, связанной с условными рефлексами и обучающимися машинами, которая давно была привнесена в литературу Винером*.

«Этим, подчеркиваю, я не хочу сказать, что условно-рефлекторный процесс осуществляется согласно указанному здесь механизму; я лишь говорю, что он *мог бы* так осуществляться.

Если, однако, принять этот или аналогичный механизм, то о нем можно сказать многое. Во-первых, этот механизм способен к обучению. Уже признано, что условный рефлекс является механизмом, способным к обучению... В природе вычислительной машины нет ничего несовместимого с наличием в ней условных рефлексов» [3].

его подмножеством $\{a\}$, содержащим этот единственный элемент. В действительности $a \not\equiv \{a\}$, а именно a есть элемент множества $\{a\}$ Какова принципиальная пропасть между понятиями множества и его элемента, иллюстрируется примером пустого множества Λ, которое как (абстрактный) объект существует, тогда как элементов, принадлежащих ему, нет.— *Прим. перев.*

* Норберт Винер (1894—1964), профессор Массачусетского технологического института, является одним из основоположников кибернетики; в течение десяти лет, предшествовавших написанию его фундаментального труда («Кибернетика, или управление и связь в животном и машине»), изучал нейрофизиологию в Национальном институте кардиологии в Мехико и сотрудничал с известным нейрофизиологом Артуро Розенблютом. Зная глубокие идеи этого труда, трудно согласиться с М. Таубе о привнесении Винером в литературу «слабой физиологии». Утверждение это остается недоказанным.— *Прим. ред.*

Когда именитые ученые начинают *приводить доводы* только на том основании, что они *могут* оказаться истинными или что *нет обстоятельств, мешающих* им быть истинными, наука становится неотличимой от предрассудков.

В задачу настоящей главы вовсе не входит отрицание возможности формирования условных рефлексов, а мы лишь приведем пример поведения, не объяснимого с точки зрения условных рефлексов. Здесь опять можно сослаться на вывод Шеррингтона о физиологическом различии между привычкой и рефлексом и на любой осознанный опыт обучения как свидетельство того, что обучение нельзя считать разновидностью условного рефлекса. Каждый, кому когда-либо приходилось обучаться игре на фортепиано, знает, что процесс обучения состоит в постепенной замене *сознательного управления* пальцами на *привычное* (бессознательное, машинальное) *движение* пальцев. Одно из приемлемых определений обучения гласит, что обучение есть замена сознательной деятельности, направленной на достижение желаемой цели, на бессознательную (привычную). Это физиологическое и психологическое определение обучения звучит так же просто и ясно, как «улучшение качества функционирования машины на основе прошлой истории ее работы». Если принять предложенное определение обучения, то машинное обучение должно предполагать существование машин, обладающих сознанием и способных вырабатывать привычки, которые могли бы некоторые сознательные реакции машин превращать в бессознательные. Я не вижу больше оснований применять это определение обучения в отношении машин, хотя первое определение, как я считаю, вполне применимо для описания обучения человека. Теперь, честности ради, не будем употреблять слова «обучение», не сделав соответствующего пояснения для разграничения обучения человека от машинного обучения.

Чрезвычайно существенным обстоятельством является тот факт, что в литературе по машинному обучению почти исключительно рассматривается обучение различным играм наподобие Тьюринговых игр, шашек или шахмат. Это весьма любопытно. Если машина может обучаться, то зачем же ограничивать ее способности обучением играм, которые, очевидно, не имеют важных социальных приложений. Ведь шахматисты и шашисты имеют меньшее зна-

чение для общества, чем футболисты. Так зачем же бросать миллионы долларов на возню с машинами, которые не приносят никакого удовольствия, не так ли?

Обычно дается следующее оправдание обучения машин игре в шахматы или шашки. Говорят, что игра в шахматы — блестящий пример бесспорно интеллектуальной деятельности человека и вместе с тем деятельности, легко выделяемой в чистом виде, что позволяет поставить ограниченный эксперимент, не засоренный неопределенными сопутствующими условиями. На первый взгляд это оправдание кажется в высшей степени разумным. И оставшаяся часть этой главы посвящена именно разбору неопределенности понятия «обучение играть в какую-то игру».

Рассмотрим сначала игру в футбол. Можно легко сообразить, что знание правил игры в футбол еще не эквивалентно способности играть в футбол. Иными словами, что касается футбола, то предложения: «Джон знает футбольные правила» и «Джон умеет играть в футбол» не эквивалентны. Здесь можно указать причину этой неэквивалентности: она кроется в том факте, что в отношении правил игры футбол представляет собой игру интеллектуальную, а с точки зрения способности играть в футбол — это игра физическая. Однако можно видеть, что то же самое различие присутствует и в играх, не требующих физической силы. Возьмите игру в покер. Ее правила очень просты. Вы можете за пятнадцать минут обучить правилам покера свою жену. С другой стороны, знать правила игры в покер и уметь *играть* в него — это две разные вещи. Бридж в противоположность покеру имеет значительно более сложную систему правил. Но опять-таки лишь после того, как человек изучит правила игры в бридж, он может начать учиться *играть* в него. Если мы перейдем от карт к шашкам и шахматам— типичным машинным играм,— то различие между знанием правил игры и умением играть в нее по-прежнему остается. Каждый может выучить правила шахматной игры за пятнадцать минут, но никто не может научиться играть в шахматы за пятнадцать минут.

В этом случае, так же как и в случае машинного перевода, нетрудно понять, что операции, которые поддаются формализации, то есть описываются с помощью набора правил, могут выполняться машиной. В той мере, в которой можно составить формальные правила для ведения

игры, включая сюда не только правила игры, но и правила оценки ее результатов, в нее может играть и машина.

С другой стороны, Клод Шеннон гарантирует, что игру в шахматы можно полностью формализовать. К сожалению, вычислительной машине, способной анализировать 1 млн. ходов в секунду, играющей в эту полностью формализованную игру, потребовалось бы 10^{95} лет для того, чтобы сделать свой первый ход. Другими словами, шахматная игра формализуется путем подчинения каждого хода команде, соответствующей правилам игры, а в партии продолжительностью 40 ходов таких возможных команд будет 10^{120}. Поскольку человек может хорошо играть в шахматы и может планировать большие последовательности ходов, то очевидно, что он играет в шахматы не формально; то же можно сказать о покере и о переводе. Отсюда следует, что разница между неформальным процессом обучения игре и формальным процессом заучивания ее правил весьма существенна.

В более широком смысле эта разница есть не что иное, как разница между обучением по инструкции или предписанию и обучением благодаря опыту. Курс на «производственное обучение» в современной системе образования служит подтверждением того, что обучение — это не формальный процесс запоминания и поведения в соответствии с некоторыми правилами.

Теперь уже нет необходимости снова повторять, что усвоение (опыт) — это сознательный процесс и обучение человека, основанное на опыте, также представляет собой сознательный процесс. Но все же необходимо ответить на следующее возражение. Могут сказать, что, дескать, человек знает, что он обучается, но не знает, «как» он это делает. «Человек может решать задачи, не зная того, как это ему удается. Этот простой факт ставит определенные условия для всех попыток осмыслить и понять процесс принятия решений человеком и решения им проблем» [4].

Трюизм, выраженный в этом и подобных ему заявлениях, состоит в том, что человек может производить то или иное действие, не умея физиологически точно объяснить, как он это делает. Он может двигать своими пальцами, не зная, каким образом нервы управляют мышцами. Он может видеть, не понимая, как кодируются сообщения и передаются из окружающей среды в мозг. Он также

может переваривать пищу, не отдавая себе отчета в том, как функционируют желудок и кишечник.

Каждый человек, интересующийся этим «как», может изучать физиологию, которая, очевидно, представляет собой нечто большее, чем то, что находится во власти специалистов по вычислительным машинам. Но даже не зная физиологии, можно наблюдать существенное различие между зрением или движением конечностей и перевариванием пищи. Видеть — значит сознавать, значит сознательно усваивать независимо от того, какова физиологическая подоплека этого процесса. Аналогично человек может двигать или не двигать руками по своей воле. Но он не сознает процесса пищеварения и не может им управлять по своему желанию. Тот факт, что человек решает задачи, во многих случаях не понимая физиологии процесса решения, может представлять интерес для метафизика, но он не имеет значения ни для человека, заинтересованного в решении данной задачи, ни даже, коли на то пошло, для физиолога. В свете этих комментариев мы можем услышать: если человек не знает физиологической сущности своих действий, то почему бы не предположить, что они напоминают действия вычислительной машины? В таком аргументе содержится столько же смысла, сколько в следующем предположении: если человек понимает принцип действия велосипеда и не знает, как устроен самолет, он может утверждать, что «самолет — это велосипед».

Тьюринг анализирует «декларативную концепцию», утверждающую, что наличие у человека сознания и отсутствие его у машины обусловливает основное различие между способностями мышления и обучения у людей и машин. Он пытается оспаривать это различие на том основании, что каждый, кто прибегает к этому аргументу, должен быть солипсистом, то есть должен считать единственно существующим лишь свое собственное сознание. Здесь трудно спорить с Тьюрингом, как трудно придумать разумный ответ на бессмысленный набор звуков. Аргумент Тьюринга оказывается таким же невероятно скудным, как и вывод Аттли относительно машин, исходящий из безнадежно путанного понимания теории множеств. Может быть, Тьюринг и творческий математик, но в философии он дилетант. Этим своим замечанием он демонстрирует распространенную форму научного заблуждения,

а именно тенденцию специалистов по вычислительной технике считать себя отцами отраслей знаний, в которых они некомпетентны. Такая некомпетентность в области физиологии давно уже дает о себе знать в дискуссиях математиков и инженеров-электриков об обучении и условных рефлексах.

В трех различных своих отрывках Бар-Хиллел проводит одну и ту же мысль: ПАВКМП (полностью автоматизированный высококачественный машинный перевод) невозможен до тех пор, пока «...не будут созданы обучающиеся машины и не будут для них составлены программы, наделяющие их разумом, подобным человеческому, знаниями и способностью приобретать знания» [5]. Такое заявление ничуть не более научно и не более содержательно, чем утверждение теолога, проповедующего, что мир не наступит до тех пор, пока Христос снова не явится на землю. Было бы более благоразумным выделять национальные денежные средства на исследования по предсказанию дня второго пришествия, чем поощрять работы по решению проблем ПАВКМП, состоящие в попытках построить обучающиеся машины, наделенные разумом, знаниями и способностью приобретать знания.

Из приведенного анализа явствует, что нашим инженерам-электрикам и энтузиастам вычислительных машин следует либо прекратить болтовню об этом, либо принять на себя серьезное обвинение в том, что они сочиняют научную фантастику с целью пощекотать читателям нервы и в погоне за легкими деньгами и дешевой популярностью.

ЛИТЕРАТУРА

1. A. M. Uttley, Conditional Probability Machines and Conditioned Reflexes, «Automata studies», ed. by C. E. Shannon and J. McCarthy, Princeton University Press, 1956, p. 255.
2. Там же, стр. 257.
3. N. Wiener, Cybernetics or Control and Communication in the Animal and the Machine, New York, John Wiley and Sons, Inc., 1948, p. 151—153.
4. A. Newell, J. C. Shaw and H. A. Simon, Chess-Playing Programs and the Problem of Complexity, «IBM J. of Research and Development», 1958, 2, No. 4, p. 320.
5. Y. Bar-Hillel, Report on the State of Machine Translation in the United States and Great Britain, Jerusalem, Hebrew University, 1959, p. 38.

5. Претензии структурной лингвистики

Граница между филологией и описательной лингвистикой, которая все четче обозначалась в течение сотен лет, недавно была полностью разрушена, и одним из главных виновников этого явился Хомский. В связи с тем, что он один из наиболее видных деятелей в новой области — области лингвистического анализа, или структурной лингвистики,— подробное изучение его работ будет уместным. Одна из его книг начинается со следующего совершенно справедливого замечания: лингвиста, занимающегося «изучением языка как некоторого орудия и почти не касающегося приложений этого орудия ...можно полностью оправдать за отказ от обращения к содержанию изучаемых им лингвистических форм» [1]. Однако в том же параграфе он указывает, что лингвист «избегает семантического аспекта при изучении лингвистических форм по той простой и достаточной причине, что, видимо, только такой подход позволит ему получить ясное представление о грамматической структуре» [2]. Хомскому следовало бы знать, что грамматика в своем традиционном смысле существенно связана с содержанием и со способами его выражения, тогда как лингвистика имеет дело с явлениями речи, интересуется их происхождением и их протеканием, и ей совершенно нет дела до нормативной грамматики. Все школьники изучают грамматику, но мало кто изучает лингвистику. Некоторые лингвисты считают синтаксис и грамматику относящимися все же к лингвистике, но они заново определяют эти понятия так, что явления, описываемые этими понятиями, имеют мало общего с традиционными синтаксисом и грамматикой. Хомский определенно интересуется грамматикой, но, как мы увидим, он понимает грамматику не как совокупность правил для адекватного выражения, а почти как физическое орудие, участвующее в действительном выражении.

Поскольку при обсуждении языкового перевода важно уяснить понятие языка, то и определение языка, данное

Хомским, заслуживает внимания: «Под языком мы будем понимать множество предложений, построенных из конечного алфавита фонем (или букв)» [3]. Человек, желающий уяснить значение некоторых понятий, находит в работе Хомского лишь болтовню о «фонемах или буквах», «морфемах или словах», причем не ясно, является здесь «или» исключающим или неисключающим. Во всяком случае, многие студенты факультетов лингвистики подтвердили бы, что:

(1) буква (графическое изображение) — это не фонема (звук), а фонема — это не буква;
(2) не существует такой вещи, как алфавит фонем, хотя алфавит из множества графических знаков может символизировать набор звуков;
(3) существует язык из фонем и язык из предложений, составленных из слов, которые построены из букв, но эти *два* языка не идентичны, короче, есть устный язык и есть язык письменный, которые не идентичны;
(4) язык (в его основном значении) состоит из звуков, а не из предложений;
(5) предложения, определяемые как упорядоченные наборы слов, начинающиеся с заглавной буквы и заканчивающиеся точкой, не могут быть сконструированы из фонем, так как нет ни заглавных фонем, ни фонетических точек, кроме как в монологах датского комедийного актера Виктора Борга;
(6) конечный набор букв может быть использован для построения множества языков, следовательно, какой-нибудь один язык (например, английский) нельзя определять только как построения из данного набора букв.

Вслед за понятием самого языка Хомский в своей работе перегружает двусмысленностью понятия «грамматики», «лингвистических структур», «лингвистических преобразований», например: «Грамматику языка можно рассматривать как *теорию* [курсив мой М. Т.] структуры этого языка» [4]. И далее: «Когда говорят об английской грамматике, то рассматривают ее как орудие, которое было создано для того, чтобы с его помощью производить английские предложения» [5]. Можно согласиться, что независимо от идеи действия, выраженного термином «производить», вполне допустимо назвать «теорию» «орудием». С другой стороны,

сказать, что грамматика создана для того, чтобы производить предложения, эквивалентно утверждению, что спряжение создано для порождения глаголов.

Но Хомский не останавливается и идет дальше этих определений грамматики. Он описывает лингвистические структуры или как «предложения-ядра», или как грамматические, то есть лингвистические, преобразования предложений-ядер. По Хомскому, «предложениями-ядрами» служат такие предложения, как «John is reading the book» и «They are flying planes» *. Очень может быть, что в природе не существует таких вещей, как предложения-ядра, а если все же они есть, то к ним правильнее было бы отнести такие выражения, как «Смотри!», «Что?!», «Тьфу!», «Нет», «Да», «Может быть». Понятие предложения-ядра Хомского будет анализироваться ниже. Но если даже допустить в качестве предложений-ядер повествовательные предложения, а также назывные и глагольные предложения и т. п., то все равно не так просто понять преобразования этих предложений.

По Хомскому, предложение «John is reading the book» можно преобразовать в предложение «Is John reading the book?», а «They are flying planes» в «Are they flying planes?». Двусмысленность второго предложения и его преобразования легко проследить при разборе неоднозначности их лингвистической структуры, то есть «They /are/ flying planes» или «They /are flying/ planes».

Хомскому следовало бы знать, что эта двусмысленность не лингвистическая, а грамматическая. Если мы хотим предложение «John is reading the book» лингвистически

---

* Перевод этих предложений на русский язык не производит эффекта, на который рассчитывал автор. Это объясняется особенностями английской грамматики, которая сообщает двусмысленность этим предложениям. Если подходить чисто формально, то возможны два перевода первого предложения в зависимости от того, считать ли is вспомогательным или смысловым глаголом: «Джон читает книгу» и «Джон есть чтение книги» и два перевода второго предложения: «Они летят на самолетах» и «Они — это летающие самолеты». В русском языке тоже нетрудно найти подобные двусмысленные предложения, например «Хочу есть глагол» или «Это стало возможным путем анализа». Вторая причина, по которой здесь не был дан русский аналог, заключается в том, что автор выбрал свои примеры с расчетом продемонстрировать смысловую неоднозначность одной из самых типичных форм английского языка: «...is reading...» В-третьих, у Хомского речь идет именно об английском языке.— *Прим. перев.*

преобразовать в вопросительную форму, мы можем сделать это, не меняя предложения, посредством изменения интонации его произношения, например «John — is reading the book?» Несомненно, изменение тона голоса, подобно изменению грамматической структуры предложения, сказывается на содержании. Аналогично на содержание влияет и выражение лица*. Каждый американец знает знаменитое изречение девы Марии: «Когда говоришь это — улыбайся». Наличие улыбки или ее отсутствие в этом случае служит главным определителем смысла, но расценивать улыбку как грамматическое средство или лингвистическое преобразование в смысле Хомского — это просто проявлять наивность при обращении к очень сложной проблеме. И если даже эту наивность нарядить марковским процессом или уравнением, то все равно этим делу не помочь.

В заключение своего анализа Хомский замечает: «Если грамматика языка должна пролить свет на то, как понимать язык, то, в частности, должно быть верно следующее: если предложение неоднозначно (его можно понимать по-разному), то его нужно подвергать сначала одному, потом другому грамматическому анализу и т. д.» [6].

Грамматика, безусловно, помогает понять язык, но не всякая смысловая неоднозначность имеет грамматическое происхождение. Существует огромное количество двусмысленностей, возникающих от недостатка контекста, например: «Он выбрал пятую» **, «Он проглядел решение».

---

*Общеизвестно, что устная речь богаче и выразительнее письменной, и это дает возможность, как правило, устно точнее и определеннее передать содержание предложения, чем в письменном виде. Тем не менее можно привести примеры обратного явления, то есть примеры того, когда устный язык не позволяет ярко и лаконично выделить определенный смысл неоднозначной формы, в то время как средствами письменного языка это сделать очень просто. На этом основаны многочисленные детские (и даже не детские) парадоксы и каламбуры. Например, неоднозначность устного предложения «вдеревневолкивидноцерковьсъели» письменно разрешается весьма просто: «В деревне «Волки» видно церковь с ели» и «В деревне волки, видно, церковь съели» и т. д.— *Прим. перев.*

** Предложение в оригинале «He took the fifth» гораздо богаче смысловой неоднозначностью вследствие чрезвычайной многозначности слова «took» и отсутствия родовых окончаний в английском языке. В следующем предложении русское слово «проглядел» удачно передает двусмысленность английского эквивалента «overlooked».— *Прим. перев.*

В одной из своих последних статей Куайн показал, что как определение фонемы, так и определение класса грамматических правил не могут быть даны формально без ссылки на семантические понятия синонимии и значимости. Результат Куайна (как и теорема Гёделя в свое время и в соответствующей связи) показывает, что попытка Хомского построить формальную грамматику и правила преобразования представляет собой попытку достичь того, что строго и логически невозможно. В некотором смысле Хомский сам понимал это, так как он предваряет свое формальное описание языка следующим замечанием: «Мы ограничимся лишь английским языком и будем предполагать, что располагаем интуитивным знанием того, что является английским предложением, а что нет» [7]. Для того чтобы понять, к чему обязывает это предупреждение, сделанное в самом начале работы Хомского, предположим, что некий трактат по логике начинается с допущения наличия интуитивной способности различать, какая комбинация символов является теоремой, а какая не является. При таких предпосылках формальный вывод становится невозможным.

Совершенно независимо от природы лингвистических преобразований позиция Хомского покоится на признании существования некой сущности, которую он называет «предложением-ядром». Лингвистическое преобразование — это операция над предложением-ядром, которая превращает его в предложение, не являющееся предложением-ядром, или, обратно, это операция, с помощью которой можно предложение, не являющееся ядром, превратить в предложение-ядро.

Отсюда ясно, что вся теория лингвистических преобразований Хомского зиждется на утверждении, что класс (конечный или бесконечный) предложений-ядер существует и может быть описан. Под словом «описан» понимается следующее: если дано конкретное предложение S, то возможно путем анализа или проверки решить, является данное предложение членом класса предложений-ядер или нет (то есть принадлежит ли S классу K). Если определение принадлежности S классу K должна выполнять машина, то это выдвигает дополнительное требование — требование формальной (то есть машинной) разрешимости проблемы истинности высказывания «S принадлежит K». Хомский не очень ясно высказывается относительно фор-

мальной разрешимости этой проблемы, однако Харрис, точку зрения которого на предложения-ядра мы еще будем изучать, выражается вполне определенно, а именно он считает машинное определение принадлежности данного предложения классу K возможным и доступным.

Требование интуитивной, эмпирической или формальной разрешимости проблемы принадлежности классу K является общим требованием, выполнение которого необходимо для того, чтобы понятие класса K имело смысл. Например, если вы заявляете, что «с помощью лингвистических преобразований любое данное предложение можно свести к пшеничному предложению», то для разрешимости этого утверждения вы должны не только уметь объяснить, что такое *пшеничный*, но также объяснить, что такое *пшеничное предложение*. Короче говоря, если вы знаете, что́ такое ядро, то это совсем не значит, что вы знаете, что такое предложение-ядро. Хомский и не пытается дать строгого определения предложений-ядер. Он описывает их мимоходом:

«Когда мы действительно проводим тщательное изучение структуры английского языка, мы обнаруживаем, что грамматику можно существенно упростить, если ядро языка ограничить небольшим множеством простых повествовательных предложений (в действительности, вероятно, конечным множеством) типа «человек ел пищу» и т. д.» [8].

И далее:

«Действительные предложения, которыми мы пользуемся в повседневной жизни, обычно не являются предложениями-ядрами, а скорее их сложными преобразованиями. Однако мы обнаруживаем, что эти преобразования, вообще говоря, сохраняют смысл неизменным, так что предложения-ядра, лежащие в основе данного предложения, можно рассматривать в некотором смысле как „элементы содержания", а с точки зрения этих элементов и следует фактически „понимать" преобразование» [9].

«Анализ преобразований, в частности, позволяет проблему объяснения, как понимать язык, частично свести к проблеме, как понимать предложения-ядра, где предложения-ядра, лежащие в основе данного предложения, следует рассматривать как элементарные строительные кирпичи, из которых посредством различных операций строится данное предложение. Так как предложения-ядра,

лежащие в основе данного выражения, по-видимому, некоторым образом охватывают основное содержание этого выражения, то такой анализ также, по-видимому, подсказывает метод изучения организации связного рассуждения — задача, которая до сих пор оставалась за пределами досягаемости лингвистического анализа» [10].

Прежде чем разбирать, что́ сказано и что́ не сказано в этих отрывках о предложениях-ядрах, сделаем несколько комментариев. Одним из грамматических приемов, повсеместно используемым специалистами, занимающимися машинным переводом, лингвистическим анализом и проектированием разумных машин, является исключение сослагательного наклонения глаголов. Это просто неверно, что Хомский провел тщательное изучение английского языка и обнаружил, что английскую грамматику можно существенно упростить. Ему следовало бы применить сослагательное наклонение: «Если бы нам пришлось провести тщательное изучение ...то мы обнаружили бы...» Тогда его читатели поняли бы, что он лишь сделал предположение, а не проводил никаких исследований, на которых якобы основаны его выводы. Заметим, между прочим, что гигантский искусственный мозг, машины-переводчики, обучающиеся машины, машины, играющие в шахматы, понимающие машины и т. п., заполонившие нашу литературу, обязаны своим «существованием» людям, пренебрегающим сослагательным наклонением. В эту игру играют так. Сначала заявляют, что, если не учитывать незначительные детали инженерного характера, машинную программу можно приравнять самой машине. Затем блок-схему программы приравнивают самой программе. И, наконец, заявление, что можно составить блок-схему несуществующей программы для несуществующей машины, означает уже существование этой машины. Точно таким путем были «созданы» машина условной вероятности Аттли, «персептрон» Розенблатта, анализатор общих проблем Симона, Шоу и Ньювела и многие другие несуществующие машины, на которые в литературе делаются ссылки как на существующие.

Допустим, Хомский действительно что-то обнаружил, а не просто предположил, что́ он обнаружил бы, если бы в действительности провел тщательное изучение структуры английского языка. Что же все-таки он обнаружил? Поскольку уместно спросить, строго ли был выполнен

математический или формальный анализ грамматики, то не менее уместным будет выяснить, что это за предложения-ядра со следующим описанием:

(1) «Очень небольшое множество... (в действительности вероятно, конечное множество)»;
(2) «вообще говоря, сохраняют смысл неизменным»;
(3) «в некотором смысле, как «элементы содержания»;
(4) «по-видимому, некоторым образом охватывают основное содержание».

Что касается пункта (1), то нечто «вероятно, конечное» является по крайней мере «возможно, бесконечным», что меняет характеристику (1) на «очень небольшое множество... (в действительности, возможно, бесконечное)». На самом деле нетрудно показать, что простые повествовательные английские предложения образуют бесконечный класс, например множество предложений типа «2 следует за 1», «3 следует за 2», «4 следует за 3» и т. д. Хомский мог бы исправить свое положение, сказав, что имел в виду только «небольшое (вероятно, конечное, но, возможно, бесконечное) множество типов предложений-ядер». К сожалению, такое исправление просто добавляет к трудности описания класса K трудность описания классов K_1, K_2, K_3, ..., K_n.

В отношении пункта (2) можно сказать, что довольно трудно понять, что́ значит «вообще говоря». Может быть, нечто, «вообще говоря, сохраняющее смысл неизменным», в частности, может менять смысл? Это не крючкотворство. Если Хомский имеет в виду, что смысл предложений-ядер инвариантен относительно их преобразований, то ему нужно было прямо так и сказать, а ничего не выдумывать. На самом же деле он так не говорит из-за очевидной ложности этого утверждения, которую сразу установит внимательный читатель.

С пунктом (3) разобраться несколько труднее. Он может означать, что «в некотором другом смысле предложения-ядра не являются элементами содержания» или «в некотором смысле элементами содержания являются не предложения-ядра, а нечто другое». Какую из этих интерпретаций ни взять — все равно очевидно отсутствие ясного описания предложения-ядра.

Наконец, пункт (4), по-видимому, некоторым образом о чем-то говорит, но о чем именно он, по-видимому, говорит, некоторым образом не совсем ясно.

Несомненно, вопрос о языке труден, и осторожного автора можно оправдать за пользование такими фразами, как «по-видимому», «некоторым образом», «вообще говоря», «в некотором смысле», «вероятно, конечный», в *рассуждениях* о языке или даже в обычном описании грамматики языка. Но подобным выражениям нет места в формальной теории грамматики и в определении лингвистических преобразований, ибо они лишают строгости определение такого коренного понятия, как предложение-ядро.

Теперь можно сказать, что Хомский не дает определения предложения-ядра, потому что это понятие хорошо известно из логики и лингвистического анализа, точно так же он не определяет преобразования, ссылаясь на то, что математики знают, что́ это такое. К сожалению, термин «ядро» встречается лишь у Харриса и у Хомского и, может быть, известен изучавшим работы Харриса. Следовательно, необходимо рассмотреть, какое значение Харрис вкладывал в этот термин. Хомский и Харрис, вероятно, прибегли к слову «ядро», чтобы ввести нечто подобное атомарному предложению или атомарной формуле в логике. Таким образом, нужно выяснить, в каком смысле выражения «атомарная формула» и «атомарное предложение» используются в логике, и проверить, будут ли состоятельны работы Хомского и Харриса, если заново не определять понятия «ядра».

Не знаю, Людвиг Виттгенштейн ли ввел понятие атомарного предложения или нет, во всяком случае, он популяризировал это понятие среди логиков и логиков-позитивистов. Виттгенштейн утверждал, что все сложные предложения суть функции истинности и, следовательно, либо тавтологические, либо контрадикторные преобразования простых, то есть атомарных, предложений. В понятие атомарного предложения вкладывается эмпирическое, а не формальное содержание. (Обратите внимание на аналогию с «элементами содержания».) Атомарное предложение, по Виттгенштейну, «соглашается» с фактами, оперирует фактами или выражает факты. Оно не имеет частей, являющихся самостоятельными предложениями. И, наконец, атомарное предложение истинно, если то, о чем в нем говорится, верно, и оно ложно, если то, о чем в нем говорится, неверно. Требование к атомарным предложениям быть либо истинными, либо ложными влечет за собой требование быть «повествовательными». Вопрос, приказание или

восклицание могут оказаться короче и проще повествовательного предложения, но, так как высказывания такого типа не имеют значений истинности, их исключают из класса атомарных предложений. В работах Хомского и Харриса нет указания на то, что предложения-ядра должны иметь значения истинности. Следовательно, нет причины им быть обязательно повествовательными, как атомарные предложения Виттгенштейна. Поскольку прежде всего учитываются длина, простота предложения и примитивный характер его смысла, ни у Хомского, ни у Харриса, ни даже у Виттгенштейна не содержится ни малейшего возражения против того, чтобы восклицания, приказания и вопросы, не говоря уже о пассиве и сослагательном наклонении, не относить к предложениям-ядрам, преобразования которых дают повествовательные предложения.

В действительности же понятие атомарного предложения почти исчезло из работ по математической логике*. Его нет в указателе терминов ни в «Методах логики» Куайна, ни во «Введении в математическую логику» Чёрча. В своей «Математической логике» Куайн говорит, что есть атомарные логические формулы, но нет никаких атомарных логических предложений. Чёрч рассматривает «сингулярные» формы как формы, имеющие одну свободную переменную; из этого не следует, что сингулярная форма — это тоже атомарная формула, так как сингулярная форма с одной свободной переменной может содержать более чем одну атомарную формулу; вместе с тем не каждая атомарная логическая формула сингулярна. Например, $x \in y$ есть атомарная формула с двумя переменными. Хотя Куайн отрицает существование атомарных логических предложений, все же, может быть, существуют просто атомарные предложения; но, так как Куайна интересует логика, а не метафизика (как Виттгенштейна), он не рассматривает их.

Рудольф Карнап говорит об атомарных предложениях, которые определяет следующим образом: «Предложение, состоящее из n-местного предиката, за которым следуют n индивидных констант**, называется атомарным предло-

* Об атомарных предложениях см. Р. Л. Гудстейн, Математическая логика, Издательство иностранной литературы, 1961, стр. 25.— *Прим. перев.*

** Пусть задано множество K произвольных объектов (предметов). Обозначим через x (индивидная переменная, или предметная

жением» [11]. Это определение допускает весьма сложные атомарные предложения и требует дополнительного определения «простейших атомарных предложений», а именно предложений, состоящих из одноместного предиката с одной индивидной константой. Но, как и Куайна, Карнапа интересуют значения истинности, и он под предложением понимает повествовательное предложение, что prima facie неприемлемо для структурной лингвистики.

Нет более нужды нам продолжать изучать роль атомарных формул или атомарных предложений в логике. По-видимому, ясно, что логики стремятся найти формулы, не содержащие в качестве своих частей других формул, для того чтобы с их помощью производить квантование более сложных формул и составлять из них функции истинности. В обычном языке подобных эмпирических или формальных требований к атомарным предложениям не предъявляется.

Даже если бы такое требование существовало, ему не удовлетворяли бы предложения-ядра, потому что существенной чертой атомарной логической формулы или атомарного логического предложения является то, что атомарная формула не должна содержать формул в качестве своих частей и атомарное предложение не должно содержать предложений в качестве своих частей. Однако Харрис совершенно определенно допускает, что «ядра» могут включать в себя определения или дополнения, которые сами являются «ядрами», и что преобразования могут выполняться произвольно.

переменная) произвольный предмет этого множества. Фиксированное значение x, то есть определенный предмет множества K, называется индивидной константой (индивидуальным предметом). *Предикат*, или логическая функция, есть высказывание относительно одной (одноместный предикат, или предикат от одной переменной) или нескольких (многоместный предикат, или предикат от n переменных, где $n = 2, 3, \ldots$) индивидных переменных, рассматриваемое как функция этих переменных, принимающая два значения — «истина» и «ложь», причем каждая переменная пробегает множество K. Например, высказывание «x имеет четыре ножки», где множество K состоит из «табуретки», «стола» и «бутылки», есть одноместный предикат, принимающий значение «истина» для констант «табуретка», «стол» и значение «ложь» для константы «бутылка». Если в некотором n-местном предикате $Q(x_1, x_2, \ldots, x_n)$ индивидные переменные заменить на индивидные константы, то получим определенное высказывание. В нашем примере: «табуретка имеет четыре ножки».— *Прим. перев.*

«Если окажется, что предложения, которые мы хотели бы оставить неразложимыми, расчленяются определенными преобразованиями на различные ядра, то мы можем отказаться от этих преобразований, а правильное применение оставшихся преобразований дает нам ядра, по размеру и типу близкие к желаемым» [12].

Далее:

«Если определение (или дополнение) содержит... слова, служащие главными членами в других ядрах, то это определение (или дополнение) преобразуется в отдельное ядро; в противном случае — не преобразуется» [13].

Признают, что существуют несомненные разногласия между Харрисом и Хомским по поводу того, могут ли предложения-ядра содержать в качестве своих частей другие ядра или потенциальные ядра (то есть части, преобразуемые посредством определенных преобразований в ядра). Но это означает, что ни того ни другого нельзя признать правым без подробного описания предложения-ядра и что при отсутствии такого описания их выводы беспредметны.

Есть еще один вопрос, который нам придется рассмотреть, прежде чем отбросить предложения-ядра и лингвистические преобразования как научные заблуждения, подобные влиянию планет в астрологии или значению шишек на черепе в френологии. Не исключено, что возможно построить структуру лингвистических операций на неопределенном первичном понятии «предложения-ядра». Можно провести параллель между этой точкой зрения и арифметикой, в которой операции возможны, даже если природа чисел остается неустановленной; а также логикой, которую можно развить даже в том случае, когда природа высказываний довольно сомнительна. К сожалению, такую параллель проводить нельзя, поскольку в арифметике и логике высказываний рассматриваются отношения между числами и высказываниями соответственно, а внутренняя природа чисел и высказываний не затрагивается. Наоборот, лингвистические преобразования связаны не с отношениями между неопределенными ядрами, а с внутренним изменением самих ядер. Другими словами, анализ преобразований необходимо опирается на теорию структуры ядер, тогда как исчисление высказываний не требует знания теории внутренней структуры высказываний. Очевидно, Харрис проанализировал эту параллель

и не стал за нее цепляться: «Это (преобразования одного и того же ядра) ... не то же самое, что сведе́ние английских предложений к их логическим эквивалентам. Аппарата логики недостаточно для выражения научных утверждений и проблем» [14]. Что бы ни означало это заявление, его можно рассматривать как эпитафию для такой науки, как лингвистика преобразований.

Бар-Хиллел пытается спасти Хомского вместе с его«предложениями-ядрами», бросив Харриса на съедение волкам.

«Хомский, ученик Харриса и многое унаследовавший от своего учителя в части терминологии и основополагающих идей, позднее пришел к использованию этих терминов в смысле, существенно отличном от вкладываемого в них самим Харрисом. Или, точнее, у Хомского такие термины, как «преобразование» или «ядро», имеют гораздо более определенный смысл, и неясности в них не превышают обычного уровня, свойственного почти всем научным терминам, тогда как у Харриса определения этих терминов далеко не точны и смысл их Харрис вынужден выводить путем далеких и неразвитых аналогий, распространяя на эти понятия методы современной абстрактной алгебры» [15].

Как Бар-Хиллелу, так и Хомскому следует подумать о том, что в один прекрасный день их юношескому задору, их способности говорить одно, а делать другое и пользоваться «не более чем обычными неясными терминами» в логических и математических рассуждениях придет конец (если только он уже не пришел благодаря нашим стараниям). Появится новое поколение младотурок, которое в свою очередь облегчит свой багаж, бросив Хомского и Бар-Хиллела на съедение волкам.

Если бы Хомский и Харрис представляли единичное явление, то количество страниц, посвященное анализу их взглядов, не было бы оправданным. Но заблуждение, которому они поддались, широко распространено и привносит мистику в занятие машинным переводом и родственные ему направления машинной обработки информации. Эту мистику можно проследить по описаниям последних работ множества организаций, опубликованным в сборнике «Текущие исследования и разработки в области научной документации», выпущенном Национальным научным советом [16]. Одна из этих организаций занимается «нормализацией выражений без существенного изменения их информационного содержания» [17]. Другая изобретает «норма-

лизованный английский язык и уже разработала [обратите внимание на форму глагола*] алгебраическое представление синтаксиса, охватывающее обширный подкласс английских предложений» [18]. Третья организация проводит эксперименты по «автоматическому анализу синтаксической структуры английских предложений» и уже заключила контракты «на исследования, посвященные анализу естественных и искусственных языков» [19]. Четвертая разработала искусственный язык, «предназначенный служить пособием в исследовательских работах ...— что-то вроде словаря» [20]. Пятая конструирует «метаязык под названием «Стандартный английский» преимущественно для описания взаимосвязанных понятий» [21]. В шестой определяется «диапазон полезных связей или функций отдельных слов или фраз», и уже сообщалось об успешном ходе «работ по автоматическому кодированию отрывков на английском или иностранном языке машинным языком (с автоматическим переводом на любой другой язык)» [22]. Седьмая организация «разработала схемы соотношений... [которые позволяют] распространить лингвистические модели Хомского на «грамматику обобщенного языка» [23]. Восьмая «занимается автоматическим программированием и методами машинного синтеза алгоритмов» [24]. Девятая рассматривает сети, «элементами которых служат семантические поля» [25]. Десятая уже недалека от создания системы, «способной автоматически преобразовывать канонические английские языковые формы соответствующие жаргонные с целью синтеза английских предложений на выходе машины-переводчика» [26]. Одиннадцатая утверждает, что новые методы анализа позволяют «разумно надеяться в далеком будущем на умение так программировать вычислительную машину, чтобы она сама училась переводить» [27]. А двенадцатая разрабатывала теорию лингвистических моделей и «могла доказать ряд важных теорем — среди них некоторые, явившиеся весьма неожиданными...» [28]. Выводы этой последней организации позволяют закончить настоящую главу следующими словами: всякие теоремы относительно лингвистических моделей могут оказаться неожиданными, чтобы не сказать обескураживающими.

---

* В оригинале употреблена форма Present Perfect, подчеркивающая завершенность, законченность действия.— *Прим. перев.*

ЛИТЕРАТУРА

1. N. Chomsky, Logical Structure in Language, «American Documentation», New York, Interscience Publishers, Inc., 1957, v. VIII, No. 4, p. 284.
2. Там же.
3. Там же.
4. N. Chomsky, Three Models for the Description of Language, «IRE Transactions on Information Theory», v. IT-2, No. 3, p. 113.
5. N. Chomsky, American Documentation, vol. VIII, No. 4, p. 286.
6. N. Chomsky, IRE Transactions on Information Theory, vol. IT-2, No. 3, p. 123.
7. Там же, стр. 114.
8. Там же, стр. 123.
9. Там же, стр. 123.
10. N. Chomsky, American Documentation, vol. VIII, No. 4, p. 291.
11. R. Carnap, Meaning and Necessity, A study in Semantics and Modal Logic, Chicago, University of Chicago Press, 1947, p. 5.
12. Z. S. Harris, Linguistic Transformations for Information Retrieval, «International Conference on Scientific Information», Area V, Preprints of Papers, Washington, 1958, p. 129.
13. Там же, стр. 130.
14. Там же, стр. 128.
15. Y. Bar-Hillel, Report on the State of Machine Translation in the United States and Great Britain, Jerusalem, Hebrew University, 1959, p. 17.
16. «Current Research and Development in Scientific Documentation», No. 5, Washington, National Science Foundation», 1959.
17. Там же, стр. 23.
18. Там же, стр. 25—26.
19. Там же, стр. 30.
20. Там же, стр. 36.
21. Там же, стр. 38.
22. Там же, стр. 43.
23. Там же, стр. 46.
24. Там же, стр. 51.
25. Там же, стр. 52.
26. Там же, стр. 55.
27. Там же, стр. 59.
28. Там же, стр. 64.

6. Отношения в системе человек — машина

Несколько лет назад Отдел военно-морских исследований начал программу работ по переоборудованию кабин самолетов под названием ANIP (Программа приборного оснащения армии и флота). Таких программ существовало много, но чем выделяется именно данная программа, так это радикальным (в буквальном смысле) подходом к проблеме взаимоотношений человека и машины. Выпущен общий отчет о ходе выполнения этой программы [1]. В данной главе будет затронут лишь один аспект этого отчета, касающийся взаимоотношений человек — машина, и в связи с этим будет разобран вопрос о состоятельности попыток создать машину, моделирующую работу человеческого мозга. Мы покажем, что истинное отношение машины к человеку есть отношение дополнения и усиления, а не моделирования.

Человеческий организм и его мозг сознательно реагируют на ограниченные спектры колебаний, приходящих из внешней среды, в которой организм функционирует. Следовательно, теоретически и практически важно сконструировать рецепторы и устройства обработки данных, способные реагировать на информацию диапазонов спектров излучения, к которым человеческий организм не чувствителен. Если этот тип информации носит устойчивый или предсказуемый характер и доставляет полную систему данных для выполнения определенных действий, можно запрограммировать машину, с тем чтобы она реагировала на такую информацию выполнением соответствующих действий. Если же информация, обработанная машиной, должна быть сгруппирована способом, который нельзя заранее предсказать, то функция машины заканчивается выдачей сведений, на которые уже должен реагировать человеческий мозг. Отношения дополнения, замещения и усиления, в которых машина находится к человеку, вовсе не являются разновидностями моделирования. Мы покажем, что последние разговоры о моделировании человеческого мозга машиной лишены интереса и не вносят никакого

плодотворного вклада в понимание взаимоотношений человека и машины.

Если конструктор проектирует машину, которая должна моделировать поведение человеческого мозга, то он должен иметь ясное представление о поведении, которое будет моделироваться, то есть самым первым требованием является требование того, чтобы каждое такое предприятие устанавливало по возможности точно и полно те виды деятельности или функции мозга, которые намереваются моделировать, имитировать или даже превзойти. К сожалению, это правда, что большинство теоретических и экспериментальных работ в этой области пренебрегает этим требованием и, следовательно, оказывается вовлеченным в классический порочный круг:

(1) предлагается конструкция машины, предназначенной для моделирования человеческого мозга, который не описан;
(2) подробно описанные характеристики машины полагаются аналогичными характеристикам мозга;
(3) затем делается «открытие», что машина ведет себя подобно мозгу; порочность [2] состоит в «открытии» того, что было постулировано.

Чтобы избежать порочного круга, необходимо начать с установления понятия видов деятельности мозга или с определения его функций. Такое определение поможет установить эмпирическую меру успешности моделирования мозга данной машиной.

Однако следут признать, что не так-то просто описать структуру или функцию мозга таким образом, чтобы установить норму деятельности, которую должна достичь или превзойти машина. И именно здесь мы обратимся за помощью к некоторым основным концепциям программы ANIP, которые со значительными упрощениями можно передать следующим описанием. За первые двадцать с лишним лет своей жизни летчик растет и развивается в околоземной среде. Органы, которые передают сообщения его мозгу, и сам мозг биологически приспособлены к функционированию именно в этой среде. Но гораздо более существен тот факт, что эти органы эволюционировали в течение миллионов лет в процессе естественного отбора и приспособились выполнять определенные действия, обеспечивающие выживание человека в его естественном окружении.

Если теперь нашего летчика поднять на высоту 10 *км* и придать ему скорость 1000 *км/час* или больше, он окажется в незнакомой и биологически неподходящей среде. Его органы чувств больше не реагируют на важные стимулы, а то, что они воспринимают, не сообщает или сообщает мало полезной информации его мозгу. В таких условиях летчика прежде всего вооружают приборами. Но, по мере того как количество и разнообразие этих приборов растет и летчику приходится все увеличивать и увеличивать скорость реакции, приборное окружение также становится неудобным для правильного функционирования мозга и органов чувств. Можно попытаться исправить это положение путем соответствующей тренировки летчика, то есть изменить среду, окружающую его в первые двадцать лет жизни. Но тренировка, рассчитанная на преодоление результата нескольких миллионов лет эволюционного развития, не предвещает большого успеха.

Однако есть другой способ, а именно использование «датчиков» и «приборов» с целью создания или моделирования в кабине пилота среды, естественной для живого организма, находящегося в этой кабине.

Допустим, человек входит в темную комнату или пещеру. Если бы он мог натренироваться в ориентировке и движении, воспринимая инфракрасное излучение, акустические колебания и радиоволны, он смог бы нормально функционировать в такой среде. Но насколько проще включить свет и тем самым создать естественную среду для бодрствующего живого организма. Быть может, и возможно научить кур в темноте не спать, а нести яйца. Однако гораздо проще провести искусственное освещение в курятник, с тем чтобы куры продолжали вести свой нормальный дневной образ жизни в квазидневных условиях.

Из этого рассуждения можно вывести весьма простое и прямое определение мозга: мозг — это орган обработки информации, поступающей из относительно стабильной среды (включая сам организм), обеспечивающий успешное поведение организма по отношению к своей среде. Конкретно глаз, передающий в мозг определенную информацию, позволяет мозгу отличить открытую дверь от закрытой, с тем чтобы в первом случае пройти в дверь, а во втором остановиться перед дверью и не набить на лбу шишку.

Существенной особенностью данного выше определения мозга является то, что оно признает прагматический,

биологический характер мозга. Оно резко отличается от любого определения или описания мозга, в котором пренебрегается биологической функцией или последняя рассматривается как несущественная характеристика небиологической сущности. Например, медицинский термометр связан с биологической средой, но его не считают биологическим организмом. Вычислительная машина тоже может использоваться в биологии, но она от этого не становится живым организмом.

Это верно, что биологический организм обладает и другими свойствами — механическими, физическими, химическими, электрическими или даже логическими. Но, подчеркивая роль мозга именно как орудия адаптации и выживания организма и данного вида, мы представляем проблему «моделирования» мозга в новом свете, невыгодном избранному кругу энтузиастов вычислительной техники.

Термин «моделирование», встречающийся при рассмотрении машин, доставляет почти столько же беспокойства, что и термин «мозг». Он, очевидно, допускает некоторую степень родства*, но, что более важно, он может быть неполным или абстрактным. Например, фунт сыра можно моделировать, положив на другую чашку весов фунтовую гирю. В данном случае моделируется вес** сыра, но не вкус, не запах, не питательные свойства. На это могут возразить, что, дескать, вес является «основной» характеристикой сыра и, коль скоро он (вес) смоделирован, остальными свойствами можно пренебречь.

К сожалению, большая часть литературы по моделированию мозга написана именно в духе этого возражения. Автор выбирает одну из характеристик мозга (или иногда даже просто постулирует ее), которую в дальнейшем считает основной. Затем он указывает, что его машина может моделировать выбранную им характеристику. Резонно заподозрить, что выбор той или иной характеристики мозга для моделирования мотивирован априорным знанием того, на что способна машина, опять же удобно, хотя снова получается порочный круг.

Вас может удивить, почему авторитетные научные журналы публикуют материал такого сорта и почему им инте-

---

* Между оригиналом и моделью.— *Прим. перев.*

** Правильнее было бы сказать масса.— *Прим. перев.*

ресуются не только читатели воскресных приложений. Ответ на этот вопрос вы найдете ниже. Но прежде необходимо убедиться в чрезвычайной важности данного выше определения, подвергнув его изучению в связи с каким-нибудь формальным рассмотрением моделирования мозга. Задача состоит в том, чтобы выбрать такое формальное рассмотрение, которое, с одной стороны, было бы типичным, а с другой — отражало бы высший уровень достижений в этой области.

В 1955 году на секции электронных вычислительных машин Национальной конференции Института радиоинженеров было решено созвать симпозиум по «проектированию моделей поведения человеческого мозга». Члены комиссии представляли первый эшелон ученых, занятых в этой области, и можно считать, что на симпозиуме были затронуты наиболее важные, если не все, вопросы моделирования. Во всяком случае, мы полагаем, что симпозиум отразил современный уровень машинного моделирования мозга, и ниже будем на него ссылаться.

Первое из того, что было зафиксировано на симпозиуме,— это требование делать различие между моделированием структуры и моделированием функции [3]. Далее, коль скоро такое различие было констатировано, участники симпозиума пришли к весьма общему соглашению относительно того, что вычислительные машины нельзя считать структурно подобными человеческому мозгу и что машинное моделирование структуры мозга является малообещающим занятием. Действительно, после того как д-р Шмитт, представляющий в комитете физиологов, заметил, что, «по всей вероятности, центральная нервная система вместе со своей памятью и вычислительными функциями представляет собой статистическую систему, в которой в запоминании каждого конкретного события принимают участие миллионы клеток и те же самые клетки одновременно хранят образы миллионов других событий» [4], д-р Минский прокомментировал это замечание, сказав, что, «по-видимому, никому не удастся [построить или предложить схему модели, которая смогла бы] показать, как такая штука работает» [5]. Еще раньше д-р Эттингер заметил, что «долго еще машины будут иметь лишь функциональное сходство с живыми организмами» [6], а д-р М. Е. Марон предлагал рассматривать подобие только входов и выходов, а на машину или на мозг смотреть как на

«черные ящики», не похожие друг на друга ни структурой, ни компонентами» [7].

Правда, д-р Шмитт указывал, что некоторые специалисты в области вычислительных машин усматривали структурную аналогию между сетями проводов в электронных вычислительных машинах и нервными сетями живых организмов и очень большое значение придавали тому факту, что нервные импульсы имеют, по-видимому, «электрическую» природу. Однако Тьюринг писал:

«То, что аналитическая машина Бэббеджа была задумана как чисто механический аппарат, помогает нам избавиться от одного предрассудка. Часто придают значение тому обстоятельству, что современные цифровые машины являются электрическими устройствами и что нервная система также является таковым. Но поскольку машина Бэббеджа не была электрическим аппаратом и поскольку в известном смысле все цифровые вычислительные машины эквивалентны, становится ясно, что использование электричества в этом случае не может иметь теоретического значения» [8].

В таком случае мы, видимо, можем обойтись без «структурного» моделирования — без Франкенштейна*, роботов и целой фантасмагории искусственных людей, сделанных из искусственной протоплазмы,— и ограничиться моделированием функций. Но здесь мы наблюдаем пробел в деятельности симпозиума. Хотя каждый и признавал моделирование функций интересным и важным видом моделирования, несмотря на это, никто не предпринял никаких попыток выделить функции, которые надлежит моделировать.

Конечно, нет большого смысла утверждать, что функция мозга — это игра в шахматы или языковой перевод и что машины-шахматисты и машины-переводчики тем самым служат успешными моделями человеческого мозга. Вероятно, конструкторы машины, которые сами увлекаются шахматами или переводом, полагают, что эти виды деятельности представляют собой основные функциональные способности человека и что именно их следует моде-

---

* В фантастической повести Мэри Шелли (жены английского поэта П. Б. Шелли) «Франкенштейн, или Современный Прометей», написанной в 1818 году, студент Виктор Франкенштейн создал и оживил искусственного человека, переставшего повиноваться своему творцу.— *Прим. ред.*

лировать. Свойство человека, лежащее в основе этих видов деятельности, можно назвать мышлением. Далее, можно было бы сказать, что функцией мозга является мышление и что автоматически думающие машины моделируют функцию мозга. Но это тоже чушь! Мозг мыслит не более, чем легкие дышат: это человек мыслит и человек дышит. Больше того, в том же смысле, к каком мозг мыслит, он осознает, желает, чувствует, воображает, ненавидит, любит и т. д. Не должна ли модель мозга выполнять также и эти функции?

Короче, определить мозг через функцию «мышления» — это значит совсем не продвинуться вперед. Как мы распознáем мышление и каким образом узнáем, что данный мозг или данная машина мыслят? И это вовсе не придирка и не каверзный вопрос. На третьей конференции по теории информации, состоявшейся в Лондоне, один из ее участников следующим образом изложил свою позицию:

«Вопрос «Может ли машина выводить умозаключения?» мы можем отнести к более общему вопросу «Может ли машина делать то, что делает человек, когда говорят, что он мыслит?» Положительный ответ на оба вопроса, предлагаемый итальянской школой, следует рассматривать как последнюю (но не первую) из следующих четырех комбинаций: а) человек мыслит — машина мыслит; б) человек мыслит—машина не мыслит; в) человек не мыслит — машина мыслит; г) человек не мыслит — машина не мыслит» [9].

Вот вам и моделирование функции. Согласно поведенческой или позитивистской метафизике, ни человек, ни машина не мыслят: человек просто совершает движения в соответствии с входными сигналами, и, видимо, можно создать машину, способную моделировать любое описанное движение человеческого тела.

Моделирование человека, основанное на одинаковой неспособности человека и машины к мышлению,— это метафизический вздор, о котором мы уже говорили. Терпимость, проявляемая по отношению к литературе о моделировании человеческого мозга, во многом обязана нашему культурному отставанию, выражающемуся в том, что механистическая материалистическая философия XIX века ныне принимается как «здравый смысл». Если в конечном счете человек — это не более чем машина в том смысле, в котором Декарт считал машинами животных (человека

он отличал от машины по религиозным соображениям), то моделирование человеческого мозга с помощью машины можно истолковать как моделирование машины с помощью машины.

Любопытно, что много лет спустя после полного отказа от материализма* и детерминизма в физике эти догмы снова были восстановлены со ссылками на поведение человека. В конце XIX столетия работа Эрнста Генриха Геккеля «Загадка вселенной» разошлась крупным тиражом. В этой книге давалось материалистическое описание вселенной, основанное на концепциях ньютоновской физики. Развитие теоретической физики положило конец популярности и значению Геккеля. Но каждый новый успех в науке сразу находит множество популяризаторов, восклицающих: «Наконец-то, теперь загадка вселенной раскрыта. Еще вчера мы думали, что человек — это «не более чем» собрание движущихся атомов. Сегодня же мы знаем, что человек — это «не более чем» очень сложная вычислительная машина».

Мне кажется, люди, сделавшие действительный вклад в развитие науки, редко бывают «не более чем чем-то». Буш и фон Нейман — создатели современной вычислительной техники никогда не злоупотребляли аналогией между человеком и машиной, однако многие их ученики и последователи в Массачусетском технологическом институте и других учреждениях не сдерживают себя в этом отношении».

Можно видеть замечательную симметрию в том факте, что Альфред Норт Уайтхед после полусотни лет занятия математикой, логикой и теоретической физикой выступил с философией живого организма, основанной на существенно биологических категориях, и что фон Нейман, занимавшийся общей логикой автоматов, заключает: «...не исключена возможность того, что в ходе этого процесса логика вынуждена будет претерпеть метаморфозу и превратиться в неврологию в гораздо большей степени, чем неврология — в раздел логики» [10].

Конечно, желательно избежать метафизической полемики на эту тему и принять prima facie различие между машиной и живым организмом. Во всяком случае, если

---

* По-видимому, автор здесь имеет в виду вульгарный материализм.— *Прим. ред.*

специалист по вычислительной технике отказывается принять эмпирическое описание биологической функции мозга и обращается к материалистической метафизике, потому что иначе ему непонятно, каким образом машина будет моделировать эту функцию, он тем самым демонстрирует разновидность извращенной принципиальности, которая неуязвима в силу порождающей ее догматической невежественности.

Теперь можно понять, почему прагматическое определение функции мозга в биологических терминах оказывает терапевтическое действие. Моделировать функции мозга — значит создать машину, способную обрабатывать информацию, с тем чтобы обеспечить выживание биологического организма, неотъемлемой частью которого она (машина) является, не говоря уже о виде, к которому этот биологический организм относится. До тех пор пока кто-нибудь не выскажет точных предложений относительно того, как построить такую машину, имеет смысл прекратить болтовню о машинном моделировании человеческого мозга.

Теперь можно вернуться к вопросу о взаимоотношении человека и машины, с которого начиналась эта глава. Коль скоро бесплодное понятие моделирования мозга исключено, открывается возможность заняться изучением действительно важного вопроса, а именно каким образом машина может расширить функции мозга путем такой обработки и преобразования информации, которые не способен выполнять мозг без помощи машины. Иначе говоря, мы хотим выяснить, каким образом машина может преобразовать среду, в которой мозг не может функционировать, в среду, в которой он функционирует. (Интересно отметить, что слово «моделирование» впервые проникло в область вычислительной техники для обозначения машинной имитации специальных условий полета самолета, то есть среды, в которой находится пилот, например в тренажере выполнения «мертвой петли».) Возьмем хотя бы термометр. Мозг вместе с органами чувств не способен точно измерить температуру среды. Следовательно, его суждения о жаре и холоде слабы и ограничены узкими пределами и плохой разрешающей способностью. Термометр же реагирует на изменение теплового излучения изменением характера отражения световых лучей, падающих на его шкалу. Мозг может очень подробно различать эти последние изменения путем сопоставления уровня ртути с соответствующим

числом на шкале. Существенным здесь является то, что машина обрабатывает сообщения, с которыми не может оперировать мозг.

Задачу термометра (в данном примере медицинского) можно расширить путем замены чисел на соответствующие фразы: «иди на работу», «отдохни», «ляг в постель», «вызови доктора», «вызывай скорую помощь». Можно было бы даже присоединить к термометру датчики, которые преобразовывали бы его показания в конкретные действия, например вызов врача по телефону. Но если бы высокое показание термометра потребовало вакцинации, то никакое программированное поведение не помогло бы. В этом случае только человеческий мозг смог бы осмыслить всю информацию о среде с точки зрения своего опыта и предназначения человеческого организма.

Можно привести еще один простой пример. Представьте себе человека, находящегося в комнате и воспринимающего в определенных узких границах тепловые, акустические и световые волны. Конечно, кроме того, он подвержен и другим видам облучения, к которым не чувствителен. В определенное время в комнату вносят радиоприемник или, еще лучше, телевизор. Телевизор — это устройство, преобразующее вид излучения, на который человеческий организм сознательно не реагирует, в знакомые ему зрительные и звуковые сигналы, которые он может воспринимать и воспринимает. Поскольку радиопередачу нельзя слушать без помощи радиоприемника или телевизора, мы будем рассматривать ее как процесс передачи сообщений приемнику; мы будем считать, что сообщения не поступают, если приемник отсутствует, хотя в действительности эти сообщения приходят. На нас непрерывно падает излучение (к которому мы не чувствительны), приходящее со всех концов земли, от небесных тел и даже от самих стен нашей комнаты.

Теперь перед создателями машин можно поставить важную и нужную задачу с полным пониманием функции человеческого мозга и функции машин в системе человек—машина. Мозг вместе с органами чувств имеет высокую степень специализации, целью которой является выживание организма в ограниченно изменяющейся среде. Мы хотели бы дать возможность человеку нормально функционировать на большой высоте, в глубинах океана, за толстой броней танка, при громадных скоростях движения

и т. д. В этих условиях сообщения, поступающие из окружающей среды в мозг, перестают передавать ему полезную информацию, то есть информацию, исходя из которой человек обеспечивает себе выживание и выполняет свое назначение. В таких условиях человек находится в полной зависимости от машин, способных преобразовывать незнакомое в знакомое, улавливать излучение от удаленных объектов и трансформировать его в сигналы, воспринимаемые органами чувств человека, использовать для работы диапазоны излучения, находящиеся за пределами человеческого восприятия, и приводить всю эту информацию к виду, который знаком человеку и имеет для него смысл.

ЛИТЕРАТУРА

1. Douglas Aircraft Company, Inc. Engineering Department, EL Segundo, California. Integrated Instrument Development Contract No. NONR 1076 (00), Report No. ES-26040, September 13, 1955.
2. O. H. Schmitt, The Design of Machines to Simulate the Behavior of the Human Brain, «IRE National Convention» (Symposium), 1955, p. 240—255.
3. A. G. Oettinger, там же, стр. 241—242.
4. O. H. Schmitt, там же, стр. 251.
5. M. Minsky, там же.
6. A. G. Oettinger, там же, стр. 242.
7. M. E. Maron, там же, стр. 249.
8. A. M. Turing, Can a Machine Think? «The World of Mathematics», ed. by J. R. Newman, New York, Simon and Schuster, Inc., 1956, v. 4, p. 2104.
9. V. Somenzi, Information Theory, Third London Symposium, ed. by Cherry, London, Academic Press Inc., 1955, p. 227.
10. J. V. Neumann, The General and Logical Theory of Automata, «The World of Mathematics», v. 4, p. 209.

7. Взаимоотношения человека и машины в системах обороны

Отождествление человека с цифровой вычислительной машиной не только грешит плохой метафизикой и сомнительно в научном отношении, но и таит в себе опасную ошибку. Дело в том, что сравнение человека с машиной в настоящее время является не только предметом философии и чистой науки. Исходя из этого сравнения, пытаются распределять исследовательские фонды, оборонные бюджеты и планировать строительство систем обороны. Следовательно, некритическое отождествление человека с машиной может привести к катастрофе. Именно в этом свете и будет дано все последующее изложение. Будет показано, например, что исследовательские усилия, направленные на создание полностью автоматизированных систем обороны, основаны на многочисленных конкретных выводах относительно природы живых организмов и моделирования [1] этих организмов. При исследовании обоснованности этих выводов необходимо снова вернуться к теме, затронутой выше, а именно к понятию возможности как руководству к выбору направления исследований и разумности отличать возможное от невозможного при выборе цели исследования. Мы также рассмотрим два сугубо технических вопроса, которые могут показаться далекими от теоретических проблем отношений между машиной и человеком и практических задач автоматизации систем обороны. Я говорю о следующих вопросах:

(1) Что означают понятия «тождества» и «подобия» как отношения, лежащие в основе научных индуктивных методов и участвующие в формулировании общих концепций?

(2) Какова роль сознания в живых организмах?

В 1958 году журнал «Бизнес Уик» анализировал многочисленные неудачные случаи применения вычислительных машин в деловых и государственных операциях [2]. В большинстве случаев речь шла о неудачах, вызванных конкретными деталями,— порочность общей идеи больших систем сводилась к деталям реализации, к инженер-

ным трудностям. Вот почему при рассмотрении проблемы отношений между машиной и человеком необходимо выйти за рамки общих рассуждений и афористических сравнений и обратиться к подробным физиологическим описаниям человеческого поведения и восприятия. Уже говорилось о том, что огромный разрыв между обещаниями и реализацией разумных машин — это просто вопрос «трудностей инженерного характера». Представление об объеме «инженерных трудностей» можно почерпнуть в истории машинного перевода, которая, как было показано выше, представляет собой чередование головокружительных предсказаний невозможного с действительными неудачами реализации этих предсказаний. Сейчас всё больше понимают, что даже приближенный машинный перевод практически неосуществим из-за множества инженерных трудностей, так как вычислительная машина с быстродействием 10^6 или 10^7 операций в секунду не идет ни в какое сравнение с человеком, пользующимся при переводе интуитивным различением содержания текста. Аналогичные препятствия ограничивают шахматные машины анализом лишь небольшого числа ходов.

Если мы оставим в покое машинный перевод и игру шахматы и обратимся к проблеме создания систем обороны и выбора стратегий, мы обнаружим, что здесь разрыв между обещаниями полной автоматизации и их выполнением становится еще значительнее. Есть причины надеяться на сужение этого разрыва, но это сужение может произойти лишь в результате уточнения и углубления наших знаний о человеке и машине, а не вследствие наивных попыток моделировать волю и творчество человека.

При разработке и проектировании сложных систем — будь то системы вооружения, системы связи, разумные машины, системы обработки информации и т. д. — конструктор системы обязан учитывать пространственные и временные условия, в которых происходит взаимодействие оператора с машиной. Насколько важно учитывать это взаимодействие, можно судить по быстрому развитию в последнее время такой специальной дисциплины, родившейся на базе синтеза психологии и техники, как техническая психология. Оператор как-то реагирует на информацию, выдаваемую ему машиной, и эта его реакция может представлять собой процесс принятия решения и контрольные проверки, необходимые для перехода к последующим

фазам работы машины. Какова реакция оператора и какова вероятность ошибки в действиях оператора — на все эти вопросы должен ответить специалист в области технической психологии, в то время как обязанность конструктора — учесть факторы, сообщенные ему этим специалистом, при проектировании пультов управления, рукояток, кнопок и т. д.

Существование проблемы «симбиоза» человека и машины привело к возникновению практических и теоретических исследований, направленных на повышение уровня автоматизации систем и уменьшение степени вмешательства человека. По-видимому, желаемой целью является полная автоматизация процесса или системы.

Принято считать, что если не практически (допустим, в данный момент времени уровень развития техники недостаточно высок), то в принципе полностью автоматизированная система обладает бо́льшим быстродействием, она дешевле, надежнее и производительнее полуавтоматической системы, в которой обязательно участие человека. Именно поэтому не пилотируемая авиация, а управляемые ракеты составляют стратегическое оружие. Аналогично это допущение породило проекты систем обороны и нападения, в которых вычислительные машины и роботы воюют против машин и роботов. Может быть, человек существует только для того, чтобы нажать первую кнопку и наслаждаться плодами победы или испытывать горечь поражения. Но это предположение тоже легко подвергнуть сомнению, и это действительно было сделано теми, кто выдумал машины, принимающие решения и обладающие творческими способностями. Такие машины, по идее их авторов, могли бы принимать свои собственные решения относительно развязывания войны и могли бы даже научиться испытывать радость или страдания от результатов своих действий.

Это все говорится не для забавы, а чтобы добиться признания того, что, если только человек не готов уйти в отставку и полностью уступить свое место машине, автоматизация не может быть полной. На некоторой стадии работы системы — пусть хоть в самом начале или в самом конце — все же останется проблема взаимоотношений человека с машиной.

Если автоматизация не может быть полной, то где, в какой степени и когда должно происходить вмешательство человека — это все вопросы меры его участия. Затем

должны существовать конструктивные принципы, исходя из которых можно определять степень участия человека и степень автоматизации для максимизации результата и минимизации капиталовложений в любую конкретную систему.

Мы отрицаем общую точку зрения, считающую увеличение доли автоматизации и уменьшение степени участия человека в системах *всегда* желательным. Мы покажем, что эта догма не только ложна, но и что ее принятие может оказать пагубное действие на нашу национальную оборону. Кроме того, мы покажем, что эта догма порождена наивностью и предубеждениями, проявляемыми в отношении способностей машин и человеческой природы. Рассуждения о «думающих машинах» сами по себе — безвредное занятие, и ему предавались еще с тех пор, как Лукреций описал мышление в виде крошечных круглых атомов, движущихся чрезвычайно быстро и отличающихся от более тяжелых и малоподвижных атомов материи. Однако когда эти рассуждения перестают быть метафизическими упражнениями и превращаются в руководящий принцип на главном направлении создания нашей системы обороны, они заслуживают того, чтобы проанализировать их самым тщательным образом. Из этого анализа необходимо сделать серьезные и реальные выводы относительно роли человека и машины в любой системе человек — машина. Точка зрения, выдвигаемая здесь и противопоставляемая только что описанной, состоит в следующем. Существуют коренные, существенные различия между цифровой вычислительной машиной и живым организмом; машина не может моделировать, копировать, имитировать или превосходить основные характеристики живых организмов. Верно также и то, что живые организмы не могут успешно копировать, имитировать или подражать основным характеристикам машин.

Если тщательно не следить за литературой, выпускаемой некоторыми из основных лабораторий, занимающихся разработкой вычислительных машин, то, возможно, трудно понять, каким образом стала превалировать догма, утверждающая тождество человека и машины. Не только серьезные технические журналы полны статей о мыслящих, обучающихся, понимающих, принимающих решения машинах, но даже популярная пресса научилась следовать моде и не моргнув глазом расписывать всяческие чудеса. Ока-

зывается, с появлением «спутника» как самой последней практической реализации идей Бака Роджерса стало немодным считать что-либо невозможным или даже практически недостижимым. В наше время укоренилась практика принимать самые неправдоподобные разглагольствования за описание завтрашних действующих устройств, и все это преподносится под соусом освещения научных перспектив. Так, журнал «Тайм» [3], освещая работу международной конференции по «Автоматизации умственных процессов», состоявшейся в Британской национальной физической лаборатории, писал, что д-р Минский, представитель Массачусетского технологического института, «... считает эту проблему искусственно усложняемой в результате неразумного преклонения человека перед своим разумом... Д-р Минский убежден в том, что в разумности и творчестве нет ничего особенного». На той же неделе в «Нью-Йоркере» [4] было помещено интервью с д-ром Фрэнком Розенблаттом, конструктором «Персептрона»— машины, которая показывает, что «небиологическая система может изучать свойства окружающей среды разумным способом... наша машина [— это] самоорганизующаяся система, [а] это в точности то, чем является любой мозг».

И снова мы можем добавить к приведенным только что цитатам отрывок из нашей любимой книжки «Исследования автоматов».

«Обучение, воспоминания и многое другое видоизменяют поведение людей в зависимости от опыта. Люди искусно разрешают проблемы, сочиняют симфонии, создают произведения искусства, литературы, техники и преследуют разнообразные цели. Они обнаруживают в высшей степени сложное поведение ... В настоящем разделе мы опишем общий метод конструирования роботов с какими-либо установленными свойствами поведения. Они могут предназначаться для того, чтобы производить *любые* желаемые и физически возможные действия в зависимости от любых обстоятельств в прошлом и в настоящем; само собой разумеется, что всякий естественный «робот», как Смит или Джонс, не может сделать большего» [5].

Перечитывая существующую литературу в поисках причин этого необычного оптимизма в отношении возможностей цифровых вычислительных машин и подобных им устройств, вы поражаетесь малочисленности действительных доказательств этих возможностей. Большинство из

таких доказательств представляет собой образцы упражнений людей с богатым воображением в мысленной экстраполяции. Фон Нейман рассмотрел процесс мысленной экстраполяции и нашел его совершенно неоправданным ни фактами, ни с точки зрения логики: «... по-видимому, естественные автоматы по своей логической организации и структуре существенно отличаются от искусственных» [6], но это его предупреждение обычно не принимают во внимание. Широко распространенное убеждение в том, что для науки нет ничего невозможного, вероятно, повинно в постоянных требованиях самых невероятных чудес [7].

Видимо, имеет смысл показать, что многие поистине важные научные достижения доказали невозможность некоторых вещей и бесплодность некоторых направлений исследований, то есть что эти исследования не могут привести к сколько-нибудь значительным результатам. Из второго закона термодинамики следует *невозможность* создания вечного двигателя; из принципа неопределенности Гейзенберга можно вывести *невозможность* точного одновременного измерения положения и скорости электрона; теория относительности утверждает *невозможность* определения понятия одновременности событий, происходящих в различных местах пространства; согласно теореме Гёделя, *невозможно* составить полную и непротиворечивую систему аксиом арифметики. Теорема Гёделя имеет прямое отношение к ограничениям возможностей вычислительных машин*, что признавали Тьюринг, Чёрч, Нагель и др. И тем не менее эти ограничения, как и указанное фон Нейманом, обычно стараются не замечать.

Конечно, понятие невозможности, если его правильно применять, в своем употреблении должно быть ограничено формальными системами, в которых его можно приравнять противоречию. Невозможности, о которых говорилось выше, формальны в том смысле, что они выводимы из общепринятых утверждений и определений, а не являются просто результатами чисто эмпирических наблюдений. В то же время множество рассуждений о *возможностях* машин

* А. Тьюринг показал, что теорема Гёделя накладывает ограничения на возможности не только машины, но и человека и в этом смысле не может служить доказательством принципиальной невозможности воспроизведения машиной операций мышления (см. А. Тьюринг, Может ли машина мыслить? Физматгиз, 1960, стр. 35—37).— *Прим. ред.*

носит чисто семантический характер. Мы уже видели, как это делается: «обучение» определяется как некоторая последовательность машинных операций и затем утверждается, что машина способна «обучаться». Подобно этому подмена доказательств декларациями стала обычным явлением при обсуждении и «обнаружении» аналогичных компонент и функций у живых организмов и вычислительных машин. Но принятие без доказательств некоторых основных определений, свойств и функций машин совершенно не оправдывает или сводит к нулю много крупных научно-исследовательских программ. Среди таких основных определений можно найти обманчиво простые (в действительности же чрезвычайно трудные) понятия «тождества» и «подобия».

Можно объяснить «подобие», исходя из понятия «тождества», и, наоборот, можно объяснить «тождество» в терминах «подобия». В зависимости от того, какое из этих двух понятий принимается за основное и неопределимое — вопрос, решаемый в большинстве случаев произвольно,— можно прийти либо к отождествлению работы машины с процессом мышления, либо к коренному различию между ними. Если «подобие» определяется в терминах «тождества и различия», то представляется удобный случай отождествить человека с машиной; с другой стороны, если понятие «тождества» определить через понятие «видимого сходства», то между человеком и машиной устанавливается коренное различие.

Допустим, «подобие» определено в терминах «тождества», которое рассматривают как известное и неопределимое понятие следующим образом: «A подобно B» означает по отношению к некоторому набору свойств $P_1, P_2, \ldots, P_n$, что

(1) некоторые элементы набора $P_1, \ldots, P_n$, а именно P_1, P_2, P_3, являются свойствами A и не являются свойствами B;
(2) некоторые элементы набора $P_1, \ldots, P_n$, а именно P_4, P_5, P_6, являются свойствами B и не являются свойствами A;
(3) некоторые элементы набора $P_1, \ldots, P_n$, а именно P_7, P_8, P_9, являются как свойствами A, так и свойствами B.

Если A — это различимая физическая сущность, которая не включает в себя B и сама не входит в B, то выделить

у A и у B идентичные свойства оказывается довольно трудным делом. Если A — это коричневый стул, а B — коричневый стол, то коричневый цвет стула не обязан быть идентичным коричневому цвету стола. Мы здесь, по-видимому, пришли к номиналистической альтернативе, в силу которой мы либо обязаны рассматривать слово «коричневый» как общий термин, применимый к двум неодинаковым окраскам,— тогда всякое неоговоренное отношение тождества исчезает,— либо мы должны принять традиционное реалистическое решение считать термин «коричневый» универсалией, то есть свойством, идентичным самому себе и независимым от того, где и когда бы оно ни проявлялось*. В этом случае коричневый цвет стола и коричневый цвет стула могли бы считаться тождественными. Понятие тождества, по-видимому, предполагает самотождественность, то есть что-то тождественно только самому себе и ничему другому. Если подобие двух объектов определено в терминах их тождества и различия, то должно существовать самотождественное свойство (или свойства), которое (которые) поделено (поделены) между этими двумя объектами.

В отчете об устройстве, названном «Персептроном» и предназначенном для распознавания подобия образов или стимулов, «подобие» определено следующим образом:

«Рассмотрим систему, экспозиция зрительного образа перед которой пробуждает активность некоторой группы клеток, которые могут быть разбросаны по всей нервной системе. Если второй образ активизирует в точности те же самые клетки, он будет считаться тождественным первому.

...Если совокупность возбужденных клеток** содержит некоторые клетки первого набора, эти два образа будут считаться «подобными», причем степень подобия пропорциональна числу возбужденных клеток, общих для обоих стимулов» [8].

В этом отрывке «подобие» определено как степень тождественности, а тождественность — это самотождественность, то есть стимулирование тех же самых клеток. В это определение трудно вложить физический смысл. Например, невероятно, чтобы клетки оставались самотождественными и неизменными в течение времени, то есть под

* Более подробно об этом см. в гл. 8.— *Прим. ред.*

** Совокупность клеток, возбужденных при экспозиции второго образа.— *Прим. перев.*

воздействием потока стимулов. И, разумеется, не существует эмпирического доказательства того, что видимое сходство цветов и звуков предполагает возбуждение пересекающихся совокупностей клеток нервной системы или мозга. Однако это определение можно интерпретировать произвольно, и, таким образом, можно оправдать конструктора «Персептрона», указав, что́ он имел в виду, употребляя слова «подобие» и «тождество». Следует также признать, что, исходя из этого определения, можно обосновать возможность формализации (и, следовательно, автоматизации) всякой задачи распознавания подобия.

Сведе́ние любой операции к программе вычислительной машины можно понимать как сведе́ние к системе совпадений и различий. Само понятие двоичного кодирования предполагает сведе́ние терминов, понятий, операций и т. п. к системе совпадений и различий. Все пробивки на перфокарте тождественны. Конечно, с физической точки зрения каждую пробивку можно отличить от другой, но любая данная пробивка по отношению к считывающему устройству считается тождественной любой другой пробивке, аналогично расположенной. Реакция считывающей головки на пробивку предполагается самотождественной в течение всего времени работы машины.

Таким образом, если допустить возможность полного разложения любого биологического или человеческого суждения относительно подобия на систему совпадений и различий, то почему бы не допустить возможность создания машины и программы решения задачи распознавания подобия для нее.

Теперь рассмотрим обратное определение, то есть примем понятие «подобия» за основное и неопределимое. Тогда тождественность объектов A и B определится как подобие свойств A свойствам B. Отсюда вытекает любопытный факт. В то время как человек распознает подобие просто, непосредственно и, по существу, творчески — как, например, мы замечаем сходство между картиной импрессиониста, поэзией Малларме и музыкой Дебюсси,— никто не смог бы подать ни малейшей идеи, как составлять машинную программу распознавания подобия; здесь уж мы не говорим об эстетической стороне творческого распознавания подобия. Такие действия постоянно встречаются в решениях военачальника, политического деятеля, коммерсанта.

Если утверждения, высказанные в отчете о «Персептроне» относительно тождественности или подобия стимулов, возбуждающих одни и те же или пересекающиеся множества клеток, рассматривать как эмпирические заявления, а не определения, то разумно рассмотреть их истинность или ложность. Как эмпирические заявления они должны устанавливать действительную связь между подобием воспринимаемых стимулов и совпадением возбуждающихся при этом нервных клеток. Если эти утверждения истинны, то любые два стимула, возбуждающие непересекающиеся множества клеток, не могут считаться подобными один другому.

Существует убедительное и почти бесспорное доказательство того, что подобие стимулов еще не влечет возбуждения пересекающихся или совпадающих множеств нервных клеток. Пятьдесят лет назад Шеррингтон опубликовал отчет о замечательной серии опытов из цикла исследований бинокулярного зрения. В этих опытах на каждую сетчатку воздействовали управляемые стимулы с тем, чтобы получить в точности подобные стимулы. Эти стимулы передавались на каждую сетчатку в течение некоторого времени синхронно, а затем в течение следующего отрезка времени с различными управляемыми скоростями. Результаты этих экспериментов в функциональных или физиологических терминах установили то, что, по-видимому, являлось анатомическим фактом, а именно что не существует общих нервных путей от обеих сетчаток и что бинокулярное восприятие не предусматривает никаких пространственных связей между нервными окончаниями в коре головного мозга. Вот что пишет сам Шеррингтон:

«Таким образом, наши эксперименты показывают, что в процессе бинокулярного восприятия внешнего образа каждый монокулярный механизм независимо от другого формирует свой полный чувственный образ. Однозначность бинокулярного восприятия является результатом объединения этих монокулярных представлений. Таким образом, однозначность бинокулярного восприятия — это результат процесса синтеза, который протекает с уже сформированными представлениями, возникающими одновременно. Мы показали, что правосторонний и левосторонний визуальные образы в мозгу локализованы различно» [9].

Сорок лет спустя Шеррингтон повторил этот вывод и подчеркнул его важность.

«Образы, проектирующиеся во взаимосоответствующие точки обеих сетчаток, вызывают один образ в мозгу — единственный мысленный образ. Это часто принималось как доказательство наличия централизации управления нервными механизмами взаимосоответствующих точек обеих сетчаток... Итак, нет свидетельств того, что нервные пути от взаимосоответствующих точек обеих сетчаток — правой и левой — соединяются в мозгу в общей точке. Однако соответствующие правостороннее и левостороннее восприятия совпадают во времени. Эта одновременность и объединяет их. Вот почему нет необходимости в их пространственном объединении в мозгу» [10].

«Персептрон», по-видимому, распознает изображенные ниже два квадрата как подобные:



Если читатель посмотрит на каждый квадрат сначала одним глазом и через некоторое время — другим, он, конечно, заметит и осознает подобие своих восприятий. Однако это подобие в действительности вовсе не предполагает возбуждения общего нервного механизма или одного и того же множества клеток головного мозга под воздействием двух подобных стимулов.

Значение этих выводов Шеррингтона состоит не только в том, что они устанавливают практическую несостоятельность представления о подобии, данного в описании «Персептрона», но и в том, что они наносят сокрушительный удар по всему развитию теории нервных сетей, впервые выдвинутой в уже упоминавшейся статье Мак-Каллока и Питтса и разрабатывавшейся далее Клини, фон Нейманом Минским, Калбертсоном и другими. Согласно теории нервных сетей, все входы нервной сети связаны с единственным выходом, причем *всегда* предусматривается пространственное объединение входов. Предполагается, что входы собраны вместе в рецепторе, или синапсе, или «маленьком черном ящике». Согласно Шеррингтону,

объединение входов в синапсе — это ложная модель объединения входов в случае бинокулярного зрения. Значит, сеть Мак-Каллока — Питтса не служит правдивым представлением нервной деятельности, во всяком случае в процессе бинокулярного зрения. Вполне вероятно также, что такая пространственная связь нервных механизмов отсутствует и при интеграции звукового образа из звуковых стимулов, поступающих через каждое ухо. И, уж конечно, никакие пространственные нервные связи не участвуют в унификации зрительных и осязательных стимулов, имеющей место в том случае, когда мы приписываем зрительные и осязательные стимулы одному и тому же предмету.

Проведенный нами анализ понятий тождества и подобия, а также конкретные результаты, доставляемые физиологией, очевидно, научно доказывают, что непосредственные аналогии, проводимые между вычислительными машинами и разумными организмами, несостоятельны. Из сказанного, очевидно, следует, что любой участок в разработке системы обороны, задача которого сводится к созданию автоматов или роботов, структурно аналогичных человеческим организмам, является совершенно ненаучным и имеет весьма малые шансы на успех.

На данной стадии наших рассуждений, однако, возникает новое осложнение в отношении проблем системы человек — машина. Хотя конструктор «Персептрона» и утверждает, что «в главном организация «Персептрона» подобна организации биологической системы» [11], другие ученые, разделяющие его основные взгляды на возможности машин, как было показано в предыдущей главе, не настаивали на *структурном* подобии или тождестве, ограничившись лишь способностью машин воспроизводить любую функцию человеческого разума.

Было замечено, что колесо, не имея никакой структурной аналогии с каким-либо органом животного, тем не менее выполняет функции конечностей животного организма. Но, несмотря на это, понятие функционального подобия слишком неясное и неопределенное и, по-видимому, относится к понятию цели. Колесо функционально подобно конечностям лишь в том смысле, что и то и другое служат подобным потребностям человека. Кроме этих целей человека, у колеса нет никаких других целей и никаких функций.

Идея машины, которая служит *целям человека*, — это здравая идея, и вопрос, на который нужно ответить, — это просто вопрос о том, смогут ли машины обеспечить достижение этих целей (или некоторых из них) с большей эффективностью, надежностью и более дешево, чем это делает человек непосредственно (без привлечения машин). Хотя этот вопрос и чисто умозрительный, так как до сих пор еще не создано ни одной машины, функционально тождественной какому-либо живому организму, все же многие виды человеческой деятельности воспроизводятся устройствами различных видов. И те, кто экстраполирует то, что есть сейчас на самом деле, очевидно, считают возможным для любой данной конкретной функции живого организма построить машину, выполняющую эту функцию (опять-таки эффективнее, надежнее, дешевле и т. п. — все это подразумевается).

Доказательство или опровержение подобного заявления не такое простое дело. Эта проблема не является ни формальной, ни семантической, но тем не менее это проблема действительная. Это значит, что правильное ее решение должно опираться на эмпирическое изучение функций живых организмов и на успех проектирования и конструирования машин, способных выполнять такие функции. Даже если нашу задачу существенно сузить за счет исключения некоторых эстетических и творческих функций и ограничиться рассмотрением защитных реакций живых организмов, она все же остается чрезвычайно трудной. Рассмотрим в качестве примера такую функцию живого организма, как сознательное принятие решения. Существует весьма убедительное доказательство того, что сознание должно вносить принципиально новое, или непредсказуемое, принятие решений в поведении животного в противоположность рефлекторной его деятельности, протекающей на бессознательном уровне и, по-видимому, не требующей никаких принципиально новых решений, то есть успешно выполняемой на бессознательном уровне. Аргументацию в пользу возможности функционального замещения этой стороны человеческой деятельности машинной можно строить в двух направлениях. С одной стороны, можно сказать, что машина тоже способна совершать принципиально новые и непредсказуемые действия (не как случайные, а как целенаправленные действия, не выводимые из машинной программы). Такой подход

предполагает возможность создания целенаправленных самоуправляемых машин, принимающих сознательные решения, не предвидимые и не предугадываемые конструктором машины. Должно быть, «Персептрон» и есть такая машина, однако «Персептрон» не существует иначе, как в виде программы, составленной для другой машины, и очень трудно, если не невозможно, понять, каким образом машинная программа, представляющая собой кодовую последовательность операций и команд, может выполнять действия, для которых принципиально нельзя составить никакой системы команд и нельзя написать никакой программы.

Второй и более естественный способ аргументации в пользу функционального замещения человека машиной строится по бихевиористической* схеме отрицания участия сознания в любой действительной функции. Коль скоро сознание исключено, принятие решений становится математически предсказуемым на основе предварительно собранных данных.

Это, конечно, верно, что многие человеческие действия основываются и все больше будут основываться на вычислении математических вероятностей. Если некоторое действие предсказуемо и все его характеристики можно рассчитать на основе имеющихся измерений, то, по-видимому, разумно поручить выполнение данного действия машине и избавить от этого человека, особенно когда эти вероятности вычисляются довольно быстро и дают более высокую меру ожидания успеха, чем неколичественное интуитивное суждение.

Можно видеть, что идея *полностью* автоматизированного ведения войны и создания автоматической системы обороны опирается на несколько допущений, которые могут быть либо истинными, либо ложными, а именно: 1) непредсказуемые действия противника, то есть действия, не предопределимые с помощью программного устройства, никогда не будут иметь места; 2) поиск оптимального решения с помощью машины всегда быстрее и надежнее, чем с помощью интуитивного суждения человека; 3) творческая импровизация никогда не потребуется в оборонительных ситуациях; 4) при планировании или проведении оборонительных мероприятий не придется прибегать

* Поведенческой.— *Прим. перев.*

к обобщениям и развитию новых методов или рассматривать новые возможные действия. Таким образом, идея полностью автоматизированных военных действий — это не только предмет размышлений над тем, что возможно и что невозможно в будущем, но это также вопрос, зависящий от возможных истинности или ложности некоторых допущений, сделанных в настоящее время.

К чему же придется прибегнуть для доказательства правомочности или установления необоснованности этих допущений? Военный противник — это не ракета и не вычислительная машина, а человек, направляющий ракеты и конструирующий вычислительные машины. Это его́ действия нужно предсказывать и против них обороняться нашими ракетами и вычислительными машинами. Теперь проблема формулируется следующим образом: можно ли действия противника (человека) запрограммировать и предсказать с помощью наших вычислительных машин таким образом, чтобы вмешательство человека и его суждения относительно принятия желательных контрмер нашей системой обороны оказались излишними?

В связи с тем что перечисленные выше допущения касаются главным образом поведения человека, по-видимому, ответ на наш вопрос нужно искать у физиологов, а не у конструкторов вычислительных машин и инженеров-электриков. Поскольку мы собираемся передать вычислительной машине функции человека, постольку прежде всего нам необходимо себе уяснить, что же это за функции. Без обращения к физиологам в данный момент все наши аргументы превращаются в бессмысленное словоблудие по замкнутому порочному кругу. Инженер-электрик объясняет нам, что́ его устройства способны делать. Однако он не может обоснованно утверждать, что они могут делать все, что́ способен делать человек, если только он не опишет человеческие способности независимо от своих машин. Если же он определяет способности человека в терминах способностей машины, то его следует обвинить в глупейшем способе решения спорного вопроса.

Дело будет гораздо вернее, если мы возьмем у физиолога описание функций человека, и тогда инженер-электрик расскажет нам, каким образом он собирается сконструировать машину, выполняющую в точности такие же функции.

Такие последовательные физиологи, как Шеррингтон и Экклс [12], утверждают, что сознательно-психическая, то есть немеханическая, интеграция опыта — это явление, оказывающее непредсказуемое влияние на действия человека. В этом утверждении «непредсказуемое» означает «принципиально непредсказуемое», а не просто «неизвестное в силу недостатка наших знаний». Возможно, слово «творческий», несмотря на его теологическую окраску, здесь подошло бы лучше, чем слово «непредсказуемый», поскольку последний термин иногда истолковывают исключительно как отрицательную характеристику наших знаний, в то время как свобода действия сознания есть свойство положительное. Бессмысленно считать, что если бы наших знаний было достаточно, то сознательные решения или действия можно было бы предсказывать. Правильно сказать так: мы имеем достаточно знаний для того, чтобы считать такое предположение бессмысленным. Точно так же бессмысленно покушаться на правильность принципа индетерминизма Гейзенберга, предполагая, что если бы мы имели достаточные знания, детерминизм снова вернулся бы в физику. Но мы имеем достаточно знаний для того, чтобы считать поведение электронов недетерминированным и отвергнуть детерминизм как физическую теорию. Аналогично мы имеем достаточно знаний для того, чтобы представлять себе роль сознания как привнесение значительной свободы и новизны в действия человека.

По-видимому, имеет смысл допустить, что в любой будущей оборонной ситуации необходимо отводить роль разумным человеческим существам, но по-прежнему остается проблема определения объема этой роли. В оставшейся части настоящей главы и будет сделано общее предположение относительно метода определения объема роли человеческого существа в системе обороны.

Если отбросить всякие предвзятые соображения относительно пределов, до которых автоматизация разумно допустима, то придется рассмотреть и проанализировать ряд оборонных проблем, начиная от индивидуальных действий летчика в его самолете и кончая тактическими и стратегическими операциями. Существенным шагом было бы согласиться с тем, что по крайней мере, если исходить из современного уровня развития науки и техники, некоторые действия, очевидно, могут выполняться машиной, а остальные, по-видимому, должны для своего

осуществления привлекать умозаключения и аналогичные способности человека, которыми не обладает машина. Подробный анализ этих областей деятельности должен выявить общее описание действий, доступных машинизации. Очевидно, существует, по крайней мере на начальной стадии такого анализа, сомнительный класс действий, в отношении которых трудно сказать, машинизируемы ли они. Но если общие критерии, по которым мы относим данное действие либо к классу машинизируемых, либо к сфере компетенции человека, вывести из анализа четко очерченных аспектов оборонительной ситуации, то эти критерии помогут сузить этот сомнительный класс.

Значение процедуры такого рода для планирования оборонительных мероприятий огромно. К практической разработке автоматической системы ведения войны, включая оборону и принятие контрмер, можно приступать хоть сейчас, но только в тех областях, автоматизация которых обещает действительный успех. И необходимо постоянно выявлять и автоматизировать все формальные звенья процессов принятия решений в системах обороны. Ни одно из указанных выше допущений не относится к таким* видам действий. В то же время нужно постоянно тренировать личный состав системы обороны с целью накопления им опыта в части решения неформальных проблем. И техническая психология тоже может здесь найти свое место, но не как теория человеческих реакций на показания шкал, циферблатов и т. п., а как наука о функциях всей сложной системы человек — машина.

ЛИТЕРАТУРА

1. O. H. Schmitt, The Design of Machines to Simulate the Behavior of the Human Brain, «IRE National Convention» (Symposium), 1955, p. 240—255.
2. «Business Week», June 21, 1958, p. 68—92.
3. «Time», **72**, No. 22, December 8, 1958, 73.
4. «The New Yorker», December 6, 1958, 44—45.
5. J. T. Culbertson, Some Uneconomical Robets, in «Automata Studies», ed. by C. E. Shannon and J. McCarthy, Princeton University Press, 1956, p. 110.
6. J. v. Neumann, The Computer and The Brain», New Haven, Yale University Press, 1958, p. 52.

---

* **Формальным или формализуемым.**— *Прим. перев.*

7. R. Bellman, On Heuristic Problem Solving by Simon and Newell, in «Operations Research», 6, No. 3, May-June, 1958.
8. Cornell Aeronautical Laboratory, The Perceptron: «A Theory of Statistical Separability» in «Cognitive Systems», prepared by Frank Rosenblatt, January, 1958, Contract No. NONR 2381 (00).
9. C. Sherrington, The Integrative Action of the Nervous System, New Haven, Yale University Press, 1906, 2nd ed., 1952, p. 380.
10. C. Sherrington, Man on his Nature, Cambridge, Cambridge University Press, 1953, p. 213—214.
11. F. Rosenblatt, The Design of an Intelligent Automaton, «Research Reviews», Washington, Office of Naval Research, October 1958, p. 6.
12. J. C. Eccles, The Neurophysical Basis of Mind, Oxford, Clarendon Press, 1956, p. 261—286.

8. Значение ⟨meaning⟩* как континуум

В настоящей книге мы постоянно отдавали себе отчет в существовании с точки зрения здравого смысла различия между символом как физическим явлением, то есть графическим изображением или звуком, и символом как носителем смыслового содержания, значения ⟨meaning⟩. В гл. 4 было указано на некоторые трудности, возникающие при попытке провести это различие более четко.

Очень трудно установить, каким образом, в каком смысле символ, знак, слово, предложение «имеют значение»; это до такой степени трудно, что как лингвисты, так и лексикографы»** стараются избегать говорить о смысловом содержании вообще. Первые подменяют «значение» понятием «зна́чимой последовательности морфем», а вторые — понятием синонимии. В отношении синонимии Куайн подчеркивал следующее:

«Вот в чем состоит этот маневр. Мы фиксируем только один главный контекст этого скверного слова «значение», а именно «подобный по значению», и начинаем обращаться с ним, как со словом «синонимичный», избегая таким образом искать значение как непосредственную сущность. Но, даже предположив, что понятие синонимии может в конечном счете дать нам удовлетворительный критерий, мы все же с помощью этого маневра избавились от всех

---

* Здесь и на протяжении всей главы слово meaning переводится на русский язык как *значение*, *смысл*, *смысловое содержание*, что согласуется с переводом этого термина в книгах: Р. Карнап, Значение и необходимость, Издательство иностранной литературы, 1959; А. Чёрч, Введение в математическую логику, т. I, Издательство иностранной литературы, 1960, а также со смыслом употребления термина meaning Куайном (например, см. «The Journal of Philosophy», 40, 1943, p. 113—127). На всех троих Таубе ссылается ниже в этой главе. Иногда в специальной литературе встречается употребление слова «значение» (не путать с принятым здесь!) как эквивалента термина Bedeutung (другие эквиваленты: *номинат*, *денотат*), использованного Фреге (см., например, Б. В. Бирюков, Теория смысла Готлоба Фреге, в сб. «Применение логики в науке и технике», Изд. АН СССР, 1960).— *Прим. перев.*

** Специалисты по составлению словарей.— *Прим. перев.*

контекстов слова «значение», кроме контекста «подобный по значению». А имеет ли данное слово также другие контексты, которые должны интересовать лингвистов? Да. Существует по меньшей мере один такой контекст — это контекст «имеющий значение»» [1].

Возможен ли глубокий анализ контекста «подобный по значению», даже если контекст «имеющий значение» оставить неопределенным,— это вопрос, который можно временно отложить в сторону. Но рассмотрение того, что́ происходит при подмене понятия «значение» понятием «зна́чимой последовательности», поможет нам сконцентрировать внимание на понятии значения как континуума.

Куайн замечает, что класс значимых последовательностей должен содержать не только все такие последовательности, которые когда-либо были произнесены или записаны, но также и все такие последовательности, которые *могут* быть произнесены. Мы пришли, пишет он, к определению состава класса значимых предложений как условных, а не действительных объектов, то есть «возможных» в отличие от действительных.

Уайтхед в свое время подчеркивал, что одна из особенностей, отличающих действительное от возможного, состоит в том, что действительное конечно и атомистично, тогда как возможное суть континуум. В этом можно видеть первый намек на то, что не поддающаяся до сих пор решению проблема смыслового содержания сможет быть наконец освещена, если не решена. Ключом к решению служит признание того, что поскольку значение включает возможное, то оно может обладать свойствами континуума, то есть класс предложений, которые могли бы быть произнесены, разумно рассматривать как несчетный бесконечный класс.

Мы можем подойти к этой проблеме и с другого направления, рассмотрев вкратце с исторической точки зрения некоторые трудности, связанные с обсуждением понятия «значение».

В этом обзоре прежних работ множество необходимых ограничений принесено в жертву краткости. Никаких подстрочных примечаний и никаких ссылок не делается, потому что цель этого резюме не установить какую-нибудь догму и не выбрать одну среди нескольких конкурирующих догм, а показать исторические традиции,

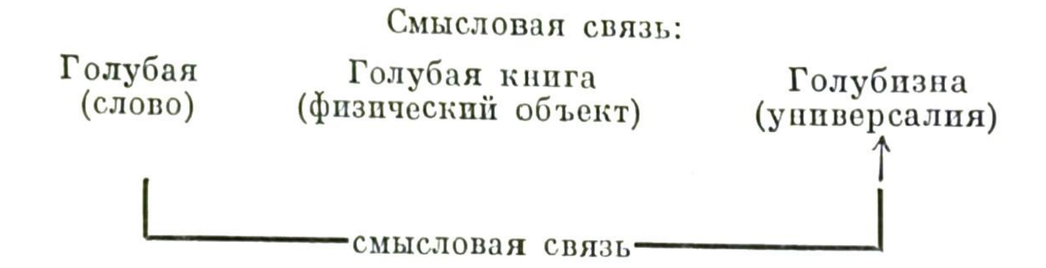
из которых сложился современный анализ значения, и предложить новый подход к этой проблеме.

Диалоги Платона в значительной своей части были посвящены изучению «значения» некоторых понятий, начиная от таких простых, как «стул», «стол», «пища», и кончая эстетическими — такими, как «справедливость», «добродетель», «красота», — и математическими — такими, как «треугольность», «прямолинейность», «точка» и т. д. Платона интересовала проблема распознавания и описания неизменной структуры значения, лежащей в основе изменяющегося потока человеческого опыта, то есть его интересовало то, что Джон Дьюи называл «Поиском определенности» ⟨The Quest for Certainty⟩. Платон утверждал, что за изменяющимся произвольным употреблением языка можно разглядеть неизменную систему «идей», или значений. Эти значения, которые плохо укладывались в рамках опыта сознания, могли бы быть раскрыты с помощью диалектического процесса чистого разума, являющегося, по существу, постепенным отходом от усвоения опыта и обращением к умозрению неизменных сущностей и значений.

Аристотель возражал против такого разделения сознания и разума (что делал Платон в некоторых своих последних диалогах) и считал идеи присущими вещам. Он указывал, что если бы стул и идея стула были двумя разными вещами, то тогда должно было существовать отношение между ними и, следовательно, идея этого отношения. Это в свою очередь должно породить отношение между отношениями вместе с идеей этого отношения и т. д. без конца. В истории философии это называется «возражением третьего члена» ⟨the third term argument⟩. И это возражение считалось сильным, так как античный мир находил неприемлемым бесконечное царство идей, или сущностей. То, о чем Аристотель смутно догадывался и что отвергал, и есть плотная несчетная бесконечность, то есть континуум.

Уайтхед заметил, что всю европейскую философию можно рассматривать как совокупность подстрочных примечаний к Платону. Одним из главных фактов, оправдывающих это замечание, служит полемика об универсалиях, пронизывающая всю средневековую философию и теологию. Что касается отношения индивидуального к универсалиям (слов к их значениям), то на этот счет существует три точки зрения, известные как реализм,

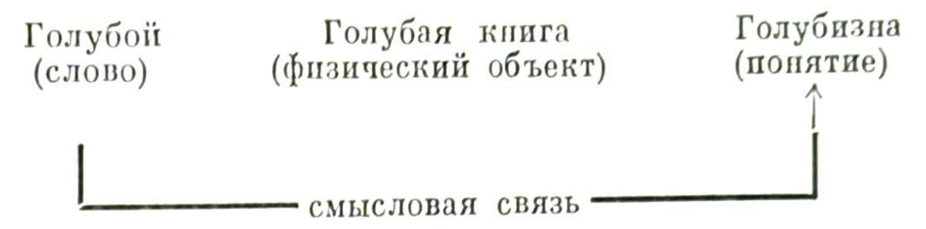
концептуализм и номинализм. Реализм явился, по существу, восстановлением платоновского мировоззрения. Мир представлялся состоящим не только из индивидуальных вещей, но и из универсалий, или сущностей. Так, голубой цвет нужно рассматривать не просто как характеристику того или иного объекта, но и как существующую независимо от последнего универсалию — «голубизну». «Голубизна» отличается от того или иного голубого предмета своей вечностью и неизменностью, то есть это универсалия. Таким образом, значение термина «голубой», рассматриваемое само по себе и взятое из любого контекста,— это просто универсалия «голубизна». Значение такой концепции состоит в том, что она описывает систему объективных идей ⟨a realm of objective meanings⟩, отличающихся как от слов, так и от естественных явлений.



По-видимому, от универсалий можно избавиться и значение слова «голубой» отнести к физической сущности, которую это слово описывает. Но поскольку голубой цвет книги — это с физической точки зрения не то же самое, что и голубой цвет чашки, то, следовательно, значение термина «голубой» изменяется так же, как изменяется его отношение к различным физическим объектам. Готлоб Фреге подчеркивал, что смысл ⟨meaning⟩ следует отличать от референта*. В следующих двух выражениях, «вечерняя звезда» и «утренняя звезда», предполагается один и тот же референт, но не один и тот же смысл. Более того, многие термины и фразы имеют смысл, если даже они не имеют отношения к физическим сущностям, например «теперешний король Франции», «$\sqrt{-2}$», «Пегас больше не полетит» и т. п. Для того чтобы избежать таких выводов и в то же время отойти от догмы реализма, была

* Подробнее об этом см. в статье Б. В. Бирюкова «Теория смысла Готлоба Фреге, в сб. «Применение логики в науке и технике».— *Прим. перев.*

разработана догма концептуализма. Концептуализм утверждает, что универсалии существуют не независимо, а как понятия (концепты) нашего разума. Понятия «голубизны», «треугольности», «справедливости», представляющие собой значения терминов «голубой», «треугольник» и «справедливый» соответственно, являются, таким образом, творениями разума; это понятия, рождающиеся в мозгу в процессе абстрагирования от физических сущностей или явлений, обладающих этими характеристиками. Таким образом, смысловая связь становится связью между словами и понятиями.



Однако концептуализм должен ответить на критику, указывающую, что это ведет к субъективизму и разрушению объективного смысла и объективной истины. Если мое понятие о «голубом» представляет собой только понятие, которое я получил в процессе абстракции от моего собственного опыта, а ваше понятие о «голубом» — это лишь понятие, полученное вами в процессе абстракции от вашего опыта, то где же гарантия тождественности моего и вашего употребления этого термина? Эту субъективность и различие в употреблении одного и того же термина нельзя считать серьезным обстоятельством по отношению к таким понятиям, как «голубой», «горячий», «холодный» и т. п.; но различия в употреблении такого термина, как «треугольник», не потерпели бы математики, а различия в употреблении таких терминов, как «добродетель», «вера», «милосердие» и т. п., не потерпела бы церковь. Концептуализм можно было спасти от обвинения в ереси только в том случае, если бы в нем нашлось место для универсального смысла ⟨universal meaning⟩. Такое место нашлось после универсализации процесса абстракции. Последнее означает, что процесс абстракции стал рассматриваться не как естественная деятельность, изменяющаяся от разума к разуму, а как рациональный процесс, одинаковый для каждого разума. Этот рациональный процесс, которым бог одарил человека, и позволяет считать, что данное понятие у одного человека тождественно соот-

ветствующему понятию у другого человека. Таким образом, отношение термина к понятию универсализировано и тем самым наука и религия спасены.

Другая точка зрения — номинализм — появилась в ходе общего наступления на разум со стороны тех средневековых теологов, которые на первый план выдвигали значение веры и мистицизма, а разум оставляли в тени. Номиналист не только отрицает существование универсалий, но вообще пытается свести понятия к именам или терминам, значение которых полностью выхолащивается и относится к конкретным объектам. «Голубой» становится словом, относящимся к тому или иному голубому объекту; «добродетель» становится словом, относящимся к тому или иному акту добродетели. Разум пользуется этими словами как знаками, но это не обязательно, говорят номиналисты и тут же вводят гипотетическую систему понятий как носителей смысла или как посредников между словами и вещами. Данное слово может относиться к множеству вещей, являясь неиндивидуальным ⟨nonspecific⟩ знаком, то есть становится общим термином, относящимся к собранию, к целой совокупности индивидуальных предметов. Такое пользование знаками не требовало никакого восприятия «общей характеристики» ⟨common characteristic⟩ и никакой выработки разумом общего понятия ⟨a general concept⟩.

Такая точка зрения приводит к полному релятивизму значений терминов не только у различных людей, но при употреблении данного термина одним человеком. Это тем не менее ни на йоту не потревожило номиналистов, а лишь нанесло удар по софистам, против которых выступал Платон, поскольку это ниспровергало претензии разума на раскрытие истины и отдавало эту область на откуп вере и откровению ⟨revelation⟩. Что же касается мнимых чисел, ложных суждений или суждений о мнимых объектах, то с их значениями поступили очень просто, отнеся их к особым идеям разума.

В современную эпоху, если не считать продолжения традиций томизма, реализм как догма почти полностью исчез до тех пор, пока его снова не возродил Чарлз Пирс у нас в Америке в середине XIX столетия и Дж. Е. Мур и Рассел в Англии в начале нашего века. Но противоречия между концептуализмом и номинализмом были пересмотрены Локком и Беркли. Локк выдвинул понятие абстракт-

ных или общих идей, порождаемых разумом на базе непосредственных идей восприятия и отражения. Он использовал понятие общей идеи, или концепта, не только для того, чтобы объяснить, каким образом разум, который начинает как tabula rasa*, развивает идею классов и свойств на основе конкретных ощущений, но также и для того, чтобы показать, каким образом разум может прийти к таким идеям, как «субстанция», «разум», «бог».

Беркли занял крайне номиналистическую позицию по той же самой причине, что и его средневековые предшественники. Он хотел дискредитировать ощущения как источник абстрактных знаний и подавить растущий материализм в науке XVIII столетия. Таким образом, он настаивал на том, что разум не в состоянии строить общие идеи на основе чувственного опыта. Мы можем употреблять термин «голубой» как слово или знак, но мы не можем иметь представления о «голубизне» как о чем-то самостоятельном, вне конкретных голубых предметов; мы не можем также представить себе «теплоту» в отрыве от конкретного проявления некоторой энергии; мы не можем представить себе форм, очертаний, не окрашенных в какой-нибудь цвет, и т. д. Что же касается таких идей, как субстанция или материя, то они были для Беркли просто общими знаками, не имеющими никакого конкретного значения или смысла.

В последнее время наметилось множество новых теорий смысла. В современных бихевиористических и прагматических теориях сделаны попытки рассматривать смысл как реакцию слушателя на выступление оратора при условии, что выступающий тоже может быть слушателем. Другими словами, смысл предложения состоит в его поведенческом эффекте. Эта точка зрения имеет специфическое следствие, состоящее в том, что данное предложение не является ни истинным, ни ложным, ни содержательным, а только станет содержательным, истинным или ложным в некоторый будущий момент. Однако в той мере, в какой эта связь с будущим является связью с возможным, должно существовать предвидение смысла как континуума, в котором отдельные значения фиксируются (или конкретизируются) конкретными предложениями.

---

* Чистая доска (*лат.*); здесь чистое место, на котором можно написать все что угодно.— *Прим. перев.*

Рассел и вслед за ним многие логики считали значение ⟨meaning⟩ свойством целых суждений, а не терминов и приписывали терминам лишь неполное значение, которое дополнялось их ролью в суждениях. Тогда следовало бы признать неполным и значение самих суждений (которое дополняется их контекстом) в том же самом смысле, в каком неполным считается значение отдельного термина. И здесь снова обнаруживается свойство значения, напоминающее континуум.

Современная точка зрения, выдвинутая Тарским и группой крайних формалистов — его последователей, явилась шагом назад благодаря введению понятия статического смыслового отношения между суждениями и фактами. Сторонники этой точки зрения надеются избежать ловушек, подстерегавших подобные теории, в которых, во-первых, предполагалось, что смысл и другие семантические понятия не могут быть формализованы средствами естественных языков, и затем строилась иерархия формальных метаязыков, каждый из которых мог содержать формальное описание семантики метаязыка, стоящего на ступень ниже его в этой иерархии. По предположению авторов, построенная таким образом иерархия метаязыков позволяет избежать гёделевской демонстрации ограничений формализации в пределах любой языковой системы.

«В метаязыке можно построить методологически правильные и существенно адекватные определения семантических понятий в том и только в том случае, если метаязык оснащен переменными более высокого логического типа, чем любая из переменных языка, являющегося предметом изучения» [2].

Поскольку нас интересует содержательная сторона естественных языков и естественных знаков, то для нас все труды Тарского оказываются неприемлемыми. Тем не менее они дают возможность продемонстрировать статическую и ограниченную природу воззрений, выдвигаемых Тарским и его последователями.

Основное в теории Тарского — это проведение строгого различия между языком, *посредством которого* мы обсуждаем нелингвистические факты, и языком, *о котором* мы говорим. Когда мы говорим о каком-либо языке, мы можем отметить этот факт посредством использования кавычек. Рассмотрим следующие предложения:

(1) Нью-Йорк — крупнейший город США.

(2) «Нью-Йорк» — это название города Нью-Йорка.

(3) Нью-Йорк содержит всего семь букв.

(4) «Нью-Йорк» содержит всего семь букв*.

Предложение (1) — это высказывание о городе. Предложение (2) — это высказывание о наименовании (города). Высказывание (3) ложно. Высказывание (4) истинно. Проведение такого различия допускает следующее определение истинности, и, по-видимому, можно дать подобные определения и для других семантических понятий.

(5) Предложение «Идет снег» истинно в том и только в том случае, если действительно идет снег. Простота этого подхода обманчива. Во-первых, Тарский отмечает, что все сказанное относится только к тому, что́ он называет классической теорией истинности, предполагающей, что истинность предложения определяется соответствием его содержания фактам. В действительности же ни одна научная система не является истинной в этом смысле. Когерентная ⟨coherence⟩ и операционная ⟨operational⟩ теории истинности, с какими бы трудностями они ни были сопряжены, безусловно, представляют более адекватное понятие научной истинности, чем эта теория соответствия. Во-вторых, пример Тарского, который, по-видимому, связывает простое предложение с простым физическим фактом, может быть принят после проверки. Но предположим, что этот его пример заменен следующим.

(6) Предложение «Бог есть» истинно в том и только в том случае, если бог действительно существует.

Не существует никакого отдельного дискретного факта
8 или ситуации, которые можно было бы сопоставить предложению «Бог есть» для установления отношения формальной эквивалентности между данным предложением и этим фактом.

Если бы Тарский ответил на это, что формальная эквивалентность выдерживается не между предложением и фактом, а между предложением в одном языке и предложением в другом, с этим можно было бы согласиться. Но тогда становится трудно понять, почему эквивалентность между предложениями в одном и том же языке — это чисто формальный факт, не имеющий семантического

* Если не считать дефиса; в английском языке слово «New York» пишется без дефиса и, таким образом, действительно содержит всего семь букв.— *Прим. перев.*

значения ⟨significance⟩, тогда как эквивалентность предложения в одном языке предложению в другом языке должна иметь такое содержание ⟨content⟩. Либо вывод Тарского чисто формален, то есть истинен по определению,— в этом случае он не имеет никакого семантического содержания; либо он должен относиться к экспериментально подтверждаемой связи между любым языком и совокупностью фактов — в этом случае существует поистине семантическое содержание, но такое, которое находится вне всякой формальной системы, даже если обратиться к бесконечной иерархии метаязыков.

Как отмечалось выше, описанная только что вкратце история теорий значения была дана без строгого соблюдения рамок и ограничений, в которых многие великие мыслители прошлого развивали свои идеи. Но такое описание было сделано с целью показать, что проблема значения прошла длинный и тернистый путь и слишком наивно искать для нее какое-нибудь простое решение, какие предлагаются новичками в этой области, считающими проблему значения новой и современной проблеме машинного перевода.

Рассматривая слово «значение», Куайн отмечает, что оно имеет по меньшей мере два контекста, именно «подобный по значению» и «имеющий значение». Но употребление Куайном термина «контекст» — это сама по себе любопытная ошибка. Очевидно, он хотел избежать употребления таких выражений, как «значение подобного по значению» или «значение имеющего значение», но при чем тут «контекст»? Нам кажется, что в обычном употреблении, а не в «контексте» выражение «подобный по значению» следует рассматривать как отношение между двумя терминами или предложениями (pMq), а выражение «имеющий значение» следует рассматривать как свойство одного термина или предложения (Ma).

Трудность такого подхода заключается в том, что он порождает проблему объяснения, какого рода отношением является «*M*». По-видимому, ясно то, что это не физическое отношение типа «больше чем», «дальше чем», «в результате» и не формальное логическое отношение, как, например, «*и*», «*или*», «*не*». Конечно, может быть, это отношение интеллектуального характера*, но поскольку еди-

* Отношение, устанавливаемое разумом.— *Прим. перев.*

ная точка зрения современного логического и лингвистического анализа, представленная в работах таких лингвистов, как Харрис, и таких логиков, как Карнап и Тарский, требует избавляться от таких отношений в пользу отношений физических и формальных, то данные отношения значения интеллектуального характера придется предать анафеме. Аналогично обстоит дело и с предикатом ⟨predicate function⟩ (*Ma*). Что же это за свойство «имеющий значение»? Это, очевидно, не физическое свойство в том смысле, в каком «красный», «тяжелый» или «длинный» являются физическими свойствами; и это не логическое свойство, как, например, свойство быть «правильно построенным» ⟨«well-formed»⟩ или быть тавтологией. И оно не может быть свойством «интеллектуального» характера — как уже было сказано, такие свойства не допускаются в формальном анализе.

Выражение «контекст» может временно избежать этих вопросов, если сослаться на то, что вопросы значения разрешимы путем анализа других слов (контекстов), среди которых оказывается слово «значение». Я думаю, что дело не в этом и что вряд ли Куайн придерживался этого мнения. Мне кажется, Куайн использовал это выражение для того, чтобы не увязнуть в вопросах, которых он не хотел затрагивать в своей статье «Проблема смысла в лингвистике», потому что сравнительно просто показать, что у него контекст означает не физический контекст, то есть начертания вправо и влево от данного слова, либо звуковые сочетания, произнесенные до или после него, а смысловой контекст. Объяснение значения через контекст подобно объяснению восприятия красного, состоящему в утверждении, что это восприятие цвета.

Если теперь мы допустим на мгновение семантическую неоднозначность двух контекстов «подобный по значению» и «имеющий значение», то сейчас самое время сказать еще о третьем контексте, а именно «являющийся значением». В логических терминах это значит, что значение не есть ни отношение, ни свойство, а индивид или переменная.

Контекст «являющийся значением» имеет важное эвристическое значение; он служит возрождению одной истины, которую незаслуженно пытаются замалчивать, истины, похороненной современными формалистами (как обычно в семье стараются отделаться от сумасшедшего кузена, предпочитая не заводить о нем разговоров). Для нас будет

вполне привычным услышать, что некоторое предложение или высказывание «имеет значение» или «подобно по значению» другому предложению. Но было бы странно слышать, что некоторое предложение «*является* значением». С другой стороны, несмотря на то, что обычно принято говорить о суждениях* как о «имеющих значение» или «подобных по значению», вполне допустимо сказать, что данные суждения (в отличие от предложений) *являются* значениями.

Подстановка терминов «предложение» или «высказывание» вместо термина «суждение» в логических рассуждениях — дело рук современных формалистов, начиная с Виттгенштейна и Карнапа и кончая Тарским, у которого формализм достиг своего апофеоза. В Principia Mathematica использованы термины «суждения», «исчисление суждений», «функции суждений». Терминов «предложения» и «функции предложений» нет в оглавлении этой книги, хотя во введении Рассела к I тому термин «предложение» случайно проскользнул вместо термина «суждение», однако это произошло лишь в неформальной части. Все формальные рассуждения касаются исключительно одних суждений, а не высказываний и не предложений. Виттгенштейн использует термин «атомарное предложение» ⟨atomic sentence⟩, а не «атомарное суждение» ⟨atomic proposition⟩. В работах Карнапа выражения «исчисление суждений» ⟨propositional calculus⟩ и «суждение» почти повсеместно вытеснены выражениями «исчисление предложений» ⟨sentential calculus⟩ и «предложение». Куайн в своей «Математической логике» рассматривает суждения. Слова «суждение» и «функция суждений» появляются в терминологическом указателе лишь со

---

* Часто в английской литературе по математической логике (в несемантических контекстах) термины sentence, statement и proposition, если между ними не предполагается различия или противопоставления, имеют общие русские значения *предложение*, *суждение*, *утверждение*, *высказывание*. Если же (в семантическом контексте) данные выражения противопоставляются одно другому или просто не отождествляются — как это имеет место здесь,— это вызывает некоторые трудности при их переводе, поскольку разные авторы вкладывают в них различное содержание. Здесь русский перевод терминов sentence (предложение), statement (высказывание), proposition (суждение) сделан в соответствии с контекстом. Вместе с тем именно такой выбор русских эквивалентов для этих терминов принят в книгах: А. Чёрч, Введение в математическую логику, т. I; Р. Карнап, Значение и необходимость. —*Прим. перев.*

ссылкой на слова «высказывание» и «матрица высказываний». Указатель к «Методам логики» Куайна уже не содержит слова «суждение», а выражение «функция суждений» появляется лишь с замечанием, объясняющим прежнее употребление этого термина. Чёрч, однако, вновь вводит в употребление термин «исчисление суждений»* в качестве названия типа логической системы, но элементами этого исчисления служат символы и формулы, а не термины и суждения. В своих «Новых основаниях» Куайн также использует слово «формула» как вместо термина «суждение», так и вместо термина «высказывание». Тарский, как уже указывалось, довел эту тенденцию до ее логического конца. Он, по-видимому, не понимает, что может существовать разница между предложением и суждением. Он рассматривает исключительно исчисление предложений и истинность как свойство предложений в формализованных языках. Однако он все же видит проблему в этой связи, о чем мимоходом замечает:

«Высказывания (предложения) рассматриваются здесь как конкретный вид выражения и, следовательно, как лингвистическая сущность. Тем не менее, если термины «выражение», «высказывание» и т. п. понимать как названия определенных последовательностей печатных знаков, то различные формулировки в этой книге оказываются не совсем корректными, появляется широко распространенная ошибка, состоящая в отождествлении выражений, одинаково изображаемых письменно ⟨of like shape⟩.... С целью избежать ...излишних усложнений в рассуждениях, связанных, кроме всего прочего, с необходимостью использования понятия подобия письменного изображения, будет удобно условиться такими терминами, как «слово», «выражение», «предложение» и т. п., обозначать не конкретные последовательности знаков, а целые классы таких последовательностей, изображаемых письменно, одинаково с данными последовательностями; только в этом смысле мы рассматриваем названия, взятые в кавычки, как индивидуальные названия выражений» [3].

Таким образом Тарский неизбежно приходит к этой проблеме. Если предложения суть физические факты, то не существует предложений, всегда повторимых.

* В нашей литературе часто как «пропозициональное исчисление» (см., например, А. Чёрч, Введение в математическую логику, т. I).— *Прим. перев.*

То, что предлагает Тарский в качестве решения проблемы, а именно — под предложением понимать целый класс подобных по начертанию последовательностей символов, — это вовсе не решение, поскольку понятие «подобия по начертанию» неформально. Сказать о знаке, изображенном на одной странице, что он есть тот же символ, что и знак, изображенный на другой странице, — значит сказать, что быть символом — это нечто большее, чем просто быть знаком определенного начертания. А это в свою очередь означает, что любой формальный процесс, как бы продолжителен он ни был и с какой бы точностью ни проводился, — это всего лишь вставка между неформальным началом и неформальным концом. Уайтхед в предисловии к «Процессу и действительности» ⟨Process and Reality⟩ требует отказаться от «веры в язык как средство адекватного выражения суждений» и категорически утверждает в первой главе, что «никакое словесное высказывание не может быть адекватным выражением суждения».

Пренебрежение к этой идее со стороны современных формалистов и то, как свободно и беспринципно они обращаются с различием между предложениями и суждениями, является одним из главных заблуждений современной науки. Не следует забывать, что как творческий математик и логик Уайтхед — это одна из величайших фигур всех времен. Великим его делает то, что в нем совмещены техническое совершенство с глубокой проницательностью. Многие из его последователей, однако, считали, что вполне достаточно сосредоточить свои усилия на технике, то есть на манипуляции символами. Вот вам пример, иллюстрирующий важность общественной миссии, которую могут выполнить научные критики. Если бы музыкант владел только техническим мастерством, не обладая ни глубиной выражения, ни интуицией, музыкальные критики разделались бы с ним как с ремесленником в лучшем случае и как с дурнем в худшем. Такие же ремесленники и дурни есть и в областях логики и математики. Но поскольку почти нет таких критиков, которые могли бы до конца понять сущность тех технических манипуляций, которыми они занимаются, те продолжают торжественно заявлять, что «техника — это все», что манипуляция символами — это самоцель, что суждениями и смыслом можно пренебречь или даже исключить их как несуществующие. В только что сказанном не сле-

дует усматривать попытку принизить значение техники; без нее и музыкант и математик остаются немыми, как бы глубока ни была их интуиция.

Что же такое суждения, которые пытаются выразить с помощью языка? Согласно Уайтхеду, суждение — это неточное понимание ⟨impure prehension⟩, принимающее участие в проявлении ⟨in the coming to be of⟩ некоторой действительной сущности ⟨actual entity⟩ или возможности ⟨occasion⟩. Каждая действительная сущность имеет физические предпосылки, физические возможности в прошлом. Кроме того, каждая физическая сущность устанавливает допустимость некоторого выбора из множества неизменных ⟨eternal⟩ объектов (сущностей), рассматриваемых как чистые возможности, которые могут быть реализованы или нет в процессе проявления этой действительной сущности. Некоторые высшие ⟨high-grade⟩ действительные сущности также формируют ⟨will experience⟩ суждения, являющиеся теоремами о допустимости некоторых возможностей в отношении определенной связи действительных сущностей. Действительная сущность связана с суждением как с орудием возможного описания этой действительной сущности через ее прошлое. Таким образом, можно сказать, что суждения играют в практике роль посредников между грубым ⟨brute⟩ фактом в прошлом и чистой возможностью в будущем. Они выражают градуированную допустимость некоторых возможностей по отношению к некоторым действительным сущностям. Понимание этой допустимости — опытное, то есть *действительное*, а не *словесное*. Каждое предложение просто предназначено для описания суждения, являющегося действительным фактом в опыте действительной сущности. С этой точки зрения суждения являются «умозрительными» ⟨mental⟩, не более чем физика, потому что это эмпирическая наука, основанная на опыте.

Это краткое описание чрезвычайно глубокого и тонкого понятия вовсе не рассчитано на то, чтобы убедить тех, кому незнакомы или кажутся странными такие слова, как «понимание» ⟨prehension⟩, «чувство, ощущение» ⟨feeling⟩, «действительная сущность» ⟨actual entity⟩ и т. п. По-видимому, есть другой способ раскрыть природу суждений, основанный на проведении различия между контекстами «имеющий значение» и «являющийся значением». Даже те, кто целиком не согласен с этим, признают,

что Уайтхед заменяет понятие «вещи» понятием «процесса» в том же самом смысле, в каком представление об атомах как о бильярдных шарах, принятое в XIX веке, позднее было заменено в физике энергетической схемой. В естественном языке «вещь» обладает свойствами, например мы говорим «этот стол красный», и если суждения — это вещи, то нужно ожидать, что суждения обладают свойствами, например смыслом, логической валентностью* и т. д. Предложение — это вещь, и предложения обладают свойствами, например они правильно построены, имеют два дюйма в длину, написаны чернилами и выражают суждения. Но суждения — это не вещи, а элементы процесса. Чтобы поставить все на свои места, Уайтхед использует выражение «пропозициональные ощущения» ⟨propositional feelings⟩. Иными словами, смысл по отношению к суждениям — это не отношение и не свойство, а индивид. Суждение — *это* смысл, а быть смыслом — значит быть элементом или сечением процесса опыта.

Теперь должны быть понятны причины таинственности и труднодоступности понятия смысла. Предложение относится к суждению как более или менее адекватное выражение смысла, которым является данное суждение. Таким образом, отношение предложения к суждению можно назвать смысловым отношением. Предложения, как и символы, обладают свойством быть выражениями суждений и как таковые обладают свойством выражать смысл. Филология и грамматика как нормативные дисциплины объясняют нам, как выражать смысл более адекватно, а лингвистика как описательная наука описывает звуки и письменные знаки, с помощью которых мы пытаемся выразить смысл. А суждение—это и есть этот самый смысл, это самое значение. Это элемент континуума значений.

Выше было показано, что в различных предлагаемых решениях проблемы значения присутствовало неявное, но настойчивое признание того, что значения не являются дискретными или отделимыми сущностями в том смысле, в котором ими являются слова. Слово, написанное на бумаге, четко отделено от других слов слева и справа от него пустыми промежутками. Предложения, написанные на бумаге, отделены от других предложений заглавной буквой (в начале) и точкой (в конце). Аналогично

* Или истинностью.— *Прим. перев.*

этому, устно произносимые слова или морфемы отделены друг от друга паузами. Довольно очевидно, что в противоположность словам и предложениям значения не отделены одно от другого ни пространственными, ни временными интервалами. Можно даже сказать, что существует не множество значений, смыслов, а одно значение, один смысл как целое пространство, делимое, но не разделенное. Суждения могут конкретизировать смысл, то есть производить сечения в пространстве смысла точно таким же образом, как физическое тело производит сечение пространства, не нарушая его непрерывности. Как бы образна ни была эта аналогия, все же многое остается неясным. Поэтому, прежде чем принимать эту теорию (теорию смысла как континуума) вместо любой или всех традиционных теорий смысла, эту туманную аналогию нужно заменить точным описанием континуума и разумным обоснованием соотнесения такого описания понятию смысла.

Мы нашли описания континуумов, применимые к смыслу, в классической монографии Хантингтона «Континуум и другие виды последовательностных построений». В этой монографии Хантингтон сначала определяет последовательность, а затем континуум как определенный вид непрерывного несчетного множества последовательностей. Для представления смысла требуется понятие неупорядоченного ⟨nonordered⟩ континуума, то есть непрерывного несчетного класса. N-мерные последовательности в смысле Хантингтона должны входить в класс значения. Однако невероятно, чтобы совокупность всех значений имела последовательностное построение ⟨serial order⟩. Хотя мы и не будем приводить здесь определения этого класса, но мы не видим причин, почему бы нельзя было сформулировать такое определение в довольно точных терминах. У нас есть более важная задача — сделать ясным и показать в более конкретных интуитивных терминах, почему смысл обладает свойствами континуума.

Прежде всего заметим, что конъюнкция или сумма суждений, есть опять суждение в том же самом смысле, в каком сумма чисел есть число и сумма пространств есть опять пространство. Не существует «самого длинного суждения», так же как не существует и «самого большого числа». Это доказывает бесконечность множества значений, но еще не доказывает его непрерывности и несчетности. Множество действительных чисел, являющееся

непрерывным и несчетным множеством, может представлять такое же множество суждений, а именно совокупность всех суждений, утверждающих существование каждого действительного числа*.

Наконец, каждое значение, которое, являясь суждением, служит выражением связи действительных сущностей посредством выбранного множества неизменных объектов, относится ⟨has relevance⟩ к несчетному множеству возможностей, представляющих собой несчетное сгущение значений, сгруппированных вокруг данного значения. Градуированная допустимость нереализованных возможностей в отношении любого действительного суждения — это как раз то, что придает суждениям различные значения у различных людей. Простое суждение «этот камень сер» — это не только позитивное понимание выражения связи действительных сущностей посредством конкретного символа, выбранного из множества возможных. Само значение этого суждения сопряжено с негативным пониманием того, чем этот камень не является. Каждому понятно, в каком смысле размер сгущения связанных допустимых значений представляет эстетический опыт. Но многие склонны пренебрегать всей этой подоплекой, когда рассматривают истинность или ложность конкретных суждений. В проведенном анализе под опытом понимался прежде всего эстетический опыт — область чувств, а не логики. Логики отвлекаются от опыта и рассматривают только определенные формальные отношения между суждениями и предложениями, с помощью которых первые выражаются. Однако и здесь мы приходим к тому же самому замечанию, с которого начали, — брать в качестве единственно возможного в действительности результат абстракции от конкретного процесса, — это основная ошибка формализма и одно из наиболее распространенных современных научных заблуждений.

Л И Т Е Р А Т У Р А

1. W. V. O. Quine, The Problem of Meaning in Linguistics, in «From a Logical Point of View, Cambridge, Harvard University Press, 1953, p. 48.
2. A. Tarski, Logic, Semantics, Metamathematics, Oxford, Clarendon Press, 1956, p. 406.
3. Там же, стр. 156.

* Это утверждение устанавливает лишь нижнюю границу мощности множества значений.— *Прим. перев.*

9. О научных заблуждениях

Настоящая книга в значительной степени обязана своим появлением книге Джона Картера и Грэхэма Полларда, вышедшей в 1934 году под безобидным названием «Исследование подлинности некоторых литературных произведений XIX века». «Исследование» Картера и Полларда представляет собой библиографический отчет о разоблачении массовых литературных подделок, опубликование которого потрясло весь ученый мир двух континентов. Авторы «Исследования», подозрение у которых зародилось в результате литературно-исторических расхождений, обнаруженных ими в предполагаемом частном издании «Португальских сонетов» английской поэтессы Элизабет Барретт Броунинг, в конечном итоге раскрыли одну из крупнейших и самых удачных литературных подделок всех времен. Лицо, совершавшее подделки, действовало успешно, продавая копии своей работы в большинство крупных собраний литературы в Америке и Англии. Ему удалось вовлечь в свои операции много выдающихся критиков и собирателей книг. Этот человек действовал по простому принципу, представляя копии своих подделок критикам, которых интересовали соответствующие авторы, а также в такие библиотеки, как Британский музей. И только после того, как восхищенные критики и библиотеки сообщали в печати о ценных экземплярах, попавших в их коллекции, автор этих подделок все же «находил» другие копии, которые продавал богатым английским и американским собирателям книг, основывая «подлинность» подделок на заявлениях Британского музея и других авторитетных критиков. Автору подделок воздавались почести за его услуги английской литературе и его коллекцию, объемистый каталог которой был куплен Британским музеем. Понятно, что раскрытие этой паутины привело бы в негодование не только самого паука, но и все солидные учреждения и литературных деятелей, попавших в его паутину.

Действительно, сама широта деятельности покрывала автора подделок. Можно было усомниться в подлинности какой-нибудь одной книги, но разве можно было осмелиться подвергнуть сомнению всю или значительную часть деятельности по сбору библиографических редкостей в Англии и Соединенных Штатах? Вот что пишут Картер и Поллард по этому поводу:

«Если мы сами были более чем удивлены результатами нашего исследования, то мы прекрасно понимаем, что столь неожиданные выводы из этих результатов не только поразят наших читателей, но и вызовут у них недоверие. Человек, всегда принимавший какую-нибудь книгу за подлинник, тем не менее готов (даже если он не заинтересован в этом как владелец книги) поверить в то, что это подделка, если доказательства покажутся ему убедительными. Но если дело касается не одной книги, а тридцати или сорока — что тогда? Мы совершенно уверены в том, что каждое из представленных нами доказательств убедительно. Но само количество обнаруженных подделок должно вызвать некоторый скептицизм у каждого, кто не подготовлен вдруг обнаружить у своих идолов глиняные ноги. Тогда скептик задаст совершенно справедливый вопрос: «Когда был совершен этот обман и кем? Каким образом все эти книги попали в хранилища? Почему этот обман оставался неразоблаченным такое продолжительное время?»

Мы в настоящей книге анализируем одно из направлений научной деятельности. Если бы мы ограничились анализом деятельности одного человека или одного учреждения, наша задача была бы куда проще. Но мы рассматриваем и высказываем свое мнение о целесообразности научных исследований, проводимых или поддерживаемых такими учреждениями, как Рокфеллеровский фонд, Массачусетский технологический институт, Гарвардский университет, РЭНД-корпорейшн, Рамо-Вулдридж и т. д. Можно было бы допустить, что любая научная деятельность, которой заняты такие организации, ipso facto* гарантирована от критики по самой своей природе. Именно в такой ситуации оказались в свое время Картер и Поллард. Тем не менее они пошли за своим исследованием туда, куда оно их привело, и их не испугал тот факт, что

---

* В силу самого факта (*лат.*).— *Прим. перев.*

результаты исследования поставили под сомнение священность священных тельцов.

Один из фактов, вскрытых Картером и Поллардом, состоял в том, что целые системы литературной критики базировались на единственном предложении, повторяемом и развиваемом различными людьми. То, что казалось внешне множеством доказательств подлинности подделок, на самом деле сводилось к следующей схеме: A цитировал B, который цитировал C, который цитировал D, который в свою очередь цитировал E и т. д. Аналогично этому в настоящей книге было показано в отношении рассматриваемой здесь научной деятельности, как то, что выглядит объемной серьезной работой, выполняемой учеными с блестящей репутацией, совершенно теряет свой блеск и убедительность, если посмотреть на это без предубеждения. В литературе по машинному переводу, машинному обучению, автоматам и т. п. можно найти множество названий несуществующих машин, работа которых рассматривается и обсуждается так, как если бы они действительно существовали. Вся эта деятельность напоминает суету портных в «Новом платье короля».

Здесь мы хотим уточнить, что параллель, проводимая между нашим исследованием и выполненным Картером и Поллардом, не распространяется на мошенничество. Мы здесь не настаиваем на том, что исследуемая нами деятельность навязывалась как преднамеренный обман, а лишь считаем эту деятельность заблуждением. С другой стороны, безотносительно к целям и намерениям ученых, попавших в сети этих заблуждений, важно подчеркнуть то, что общественность, благоговеющая перед именем ученого XX века, оказалась обманутой.

Критику науки в XX столетии можно уподобить разновидности lèse majesté*, в каком-то отношении эквивалентной критике римской католической церкви в XII столетии. Не всегда понимают, что все виды интеллектуальной деятельности, за исключением науки, имеют как профессиональных практиков, так и компетентных критиков. В области поэзии, беллетристики, драматургии, живописи, скульптуры, музыки и т. п. эти критики не только являются интерпретаторами творчества профессиональных практиков для широких кругов общества, но они

* Оскорбление величества (*фр.*).— *Прим. перев.*

также сравнивают и оценивают работу практиков. Критик, никогда не писавший стихов, партитур, картин, тем не менее может оказать обществу ценную услугу, указав на то, что данное конкретное произведение искусства (стихотворение, партитура, картина) написано без вдохновения, неестественно или подражательно. Кроме того, никто всерьез не думает, что работа таких критиков представляет собой цензуру или ограничивает свободу творчества поэтов, музыкантов или художников. А вот ученые почему-то считают любую критику их работы как «зажим» или непонимание необходимости проведения «основополагающих» (что иногда следует читать как «ненаучных» или «бесполезных») исследований. Ученые даже формально добились освобождения от всякой отчетности за расходуемые ими общественные средства, что нашло отражение в первом законе Национального научного совета, на который бывший президент Трумэн все же нашел мужество наложить вето.

Во всех общественных дискуссиях о разрыве в области ракетостроения между нами и русскими виновными в этом разрыве считают кого угодно, только не ученых. Причины наших неудач видят в слепоте администрации, в подозрительности наших военных, в политическом трюкачестве конгресса, в продажности наших корпораций или апатии общественности. Но никто никогда не сказал — а дело, по-видимому, именно в этом, — *что, возможно, наши ученые не так хороши* и что по крайней мере в этой области русские ученые более образованы и более изобретательны, чем наши.

Одна из причин такого самообмана состоит в том, что термины «ученый» и «наука» приняли в наше время оценочный оттенок и сейчас употребляются в смысле «хороший ученый» и «важная наука» соответственно. «Плохого» ученого изображают в кино как человека, жестокого по отношению к своей матери или состоящего в связи с агентами противника, тогда как в действительности плохой ученый, как и плохой поэт или плохой музыкант, — это человек, работы которого плохи. Есть плохая наука в том же самом смысле, в каком существуют плохая поэзия и плохая музыка. Существуют тысячи ученых, имеющих к науке такое же отношение, как многие люди к религии, политике, музыке, торговле, рекламной деятельности и т. п. Ученые — это обычные люди, среди которых есть

преданные своему делу и своекорыстные, умные и дураки, честные люди и проходимцы. Сегодняшние ученые признают, что существуют как просто ограниченные ученые, так и неглупые ученые, совершившие ошибку, но обычно они выносят такое суждение лишь о прошлом. Имеются в виду, по-видимому, только те научные ошибки, которые были совершены давным-давно, и для широких масс образ ученого — это образ какого-то особенного человека, который не посмел бы допустить ошибку и выбросить на ветер миллионы долларов народных денег по своей глупости или недостаточной компетентности.

Целью настоящей книги является не оценка per se* работы отдельного ученого, а общий анализ научных заблуждений сегодняшнего дня. Ученые охотно согласятся с тем, если вы скажете им, что существовали научные заблуждения в прошлом — ведь существовали же такие науки, как алхимия и френология. Признáют и то, что наряду с истинно научными направлениями в научном мире подвизаются такие псевдонауки, как астрология, хиромантия, спиритизм. Но четко ли обозначена граница между наукой и псевдонаукой? Что можно сказать относительно телепатии, парапсихологии и общей семантики? Можно пойти еще дальше и назвать здесь психоанализ, теорию прогрессивного воспитания, науку об общественном мнении, косметологию и т. д. Все поставленные здесь вопросы предназначены для того, чтобы показать, что ни признание общественности, ни высоконаучные названия не могут служить надежной гарантией подлинной научности и истинной ценности всякого научного исследования.

И все же написано много книг, в которых обсуждается, можно ли назвать наукой ту или иную область человеческой деятельности. В области психологии различные школы, например представители ассоциативной психологии, экспериментальной психологии, интроспективной психологии, бихевиористы, психоаналитики, в разное время претендовали на научный характер своих работ и отвергали претензии других школ на право называться наукой. У нас в Америке была дискуссия относительно права общественных наук называться науками и укрываться под крылом Национального научного совета.

* Самой по себе (*лат.*).— *Прим. перев.*

Те, кто занимается такой наукой, как рекламное дело, в основе которого лежит теория случайного выбора, калькуляция и учет изменчивых человеческих побуждений, учились по сути дела только для того, чтобы под ученой мантией торговать сигаретами, косметическими товарами и бутербродами. Вся беда в том, что под словом «наука» стали понимать нечто вроде обряда посвящения в рыцари. Это слово — как лавры, которыми увенчивают какое-нибудь предприятие и тем самым создают ему добрую репутацию.

Было время, когда слово «религия» носило такой же оценочный смысл. Быть религиозным означало ipso facto быть хорошим человеком; быть нерелигиозным или антирелигиозным значило ipso facto быть плохим человеком. Но потом это слово потеряло свой оценочный характер, так что в наше время принято говорить как о хорошей религии*, так и о плохой религии. Другими словами, в отношении религии признается важным не сам факт, что вы исповедуете ту или иную религию, а то, *что́* именно вы исповедуете.

Аналогичная метаморфоза термина «наука» была бы очень полезной. Тогда не будет большой беды в том, если каждый, кому так заблагорассудится, станет называть себя «ученым», а сферу своей деятельности — «наукой». Тогда можно будет прекратить спор о значении этого слова и заняться более важной задачей — задачей оценки самой деятельности. А сейчас дела обстоят таким образом, что если кто завладеет правом называть свою деятельность «наукой», то его никто не смеет назвать шарлатаном, даже если он таковым является. Возьмем астрологию и астрономию. Если астролог претендует на звание ученого так же, как и астроном, то я должен им обоим дать carte blanche** в отношении использования этого слова (слова «ученый»). А carte blanche сразу же дает право эксплуатировать магические свойства этого слова. Если бы быть ученым значило не больше, чем быть высоким или коротким, толстым или тонким, слово «наука» не использовали бы для гипнотизирования доверчивой публики.

* Советскому читателю ясно, что научное мировоззрение несовместимо ни с какой религией, в том числе и с «хорошей». Однако автор, по-видимому, под термином «религия» понимает более широкую систему взглядов.— *Прим. ред.*

** Donner à qn carte blanche (*фр.*).— предоставить кому-либо полную свободу действий.— *Прим. перев.*

Тогда деятельность любого гражданского или закрытого учреждения будет оцениваться как и всякая деятельность и ответственным за эту деятельность не избежать проверки, не уклониться от требования истинно научной природы проводимых исследований.

Сравнение таких наук, как астрономия и астрология, наталкивает на интересные размышления. Допустим, университет располагает хорошим телескопом, но бюджет кафедры астрономии настолько мал, что не позволяет полностью использовать физических возможностей этого телескопа. Предположим далее, что честолюбивый молодой профессор представляет на заседании кафедры план расширения бюджета кафедры путем продажи времени использования телескопа астрологам. Глава кафедры может возмутиться и отказать в предоставлении даже свободного времени, то есть времени, когда телескоп не используется, для такой явно антинаучной деятельности.

Тогда молодой профессор, зная, что тысячи честных газет ежедневно помещают на своих страницах гороскопы, может сказать своему шефу: «Чисто логически вполне возможно, что зодиакальное движение планет оказывает влияние на поведение и действия человека. В допущении такой возможности нет ничего противоречивого». Может быть, глава кафедры лично сам и допускает это, но, если он честный человек, он пошлет к черту молодого профессора. Что же касается самого молодого профессора, то если он на самом деле не верит в зодиакальное влияние и тем не менее на деле способствует распространению астрологической информации, он подобен мошеннику, разоблаченному Картером и Поллардом. Но если он действительно считает, что ученому вполне достаточно чисто логической возможности для оправдания соответствующей деятельности, то он просто дурак, а не мошенник.

Рассмотрим теперь аналогичный случай, но более близкий к теме нашей книги. Предположим, что некий университет приобрел большую и дорогую вычислительную машину, которая может выполнять все необходимые в течение дня вычисления за какую-то долю всего имеющегося в ее распоряжении машинного времени. Увидев, как эта большая и дорогая машина значительную часть времени простаивает без работы, молодой профессор решает, что было бы интересно заставить эту громадину переводить, моделировать мозг или играть в шахматы.

Если его начальник спросит, действительно ли машина способна все это делать, профессор может ответить, что в этих предполагаемых способностях машины нет ничего логически невозможного. Такой ответ нельзя расценивать как обман, если только профессор не знает или не уверен в том, что в конце концов машина не сможет переводить или вести себя подобно мозгу. Но если он исходит лишь из чисто логической возможности таких способностей машины, то он такой же осел, как и тот честный профессор-астролог с кафедры астрономии.

Заблуждение, с которым мы боремся в этой книге, представляет собой не такую простую и старую мишень, как астрология или алхимия. Это новая звезда на научном небосводе — кибернетика. Основатель кибернетики Норберт Винер описал ее как науку об управлении в человеке и машине. Один из основных приемов кибернетики состоит в проведении массовых аналогий между функциями управления в человеке и такими же функциями в машине. После этого делается вывод, что изучение машинных функций управления должно сделать вклад в нейрофизиологию, изучение нейрофизиологии должно оказать помощь при разработке и конструировании машин. Несомненно, аналогии между машиной и человеком существуют и изучение этих аналогий является важным и ценным интеллектуальным занятием. Однако как фон Нейман, так и многие другие подчеркивали, что есть громадная логическая и практическая разница между аналогией и тождеством. Существует глубокое различие между функционированием человека и работой машины, и превращение отношения подобия, лежащего в основе аналогии, в формальное логическое отношение тождества является серьезной логической ошибкой, которая в рассмотренном нами случае приводит к крайне неприятным последствиям*. В истории науки

* Если пренебречь некоторыми популярными статьями, написанными неспециалистами, а говорить о серьезных работах по кибернетике, то обвинение авторов последних в утверждении тождества между функционированием человека и работой машины следует признать необоснованным. Точно так же несправедливо следующее ниже обвинение в установлении тождества между нейронными сетями и «электрическими контурами» ЭВМ. В действительности речь идет о моделировании нейронных сетей и воспроизведении при помощи этих моделей отдельных черт функционирования мозга. Вообще в этой главе наиболее полно проявляется «нигилистическое» отношение М. Таубе к перспективам кибернетики, которое отмечено академиком А. И. Бергом в предисловии к настоящей книге.— *Прим. ред.*

и философии эта ошибка много раз приводила к редукционистским* методологическим заблуждениям. Например, истолкование слышимых звуков в терминах частоты и амплитуды колебаний приводило по редукции к ложному утверждению, а именно к тому, что слышимые звуки тождественны амплитуде и частоте колебаний. В кибернетике аналогии между нейронными сетями и электрическими схемами вычислительных машин приводили по редукции к ложному выводу о том, что якобы нейронные сети тождественны электрическим контурам в вычислительных машинах.

Нет необходимости приводить дальнейшие примеры. Отметим только следующий важный факт. Фон Нейман, который, по-видимому, является самым крупным ученым своего поколения в области аналогий между человеком и машиной, предсказывал незадолго до своей смерти, что, изучая нейрологию, можно узнать о машинах гораздо больше, чем узнать о нейрологии, изучая машины.

Поскольку существует тенденция смешивать как устные или письменные символы с их значениями—этой теме было посвящено несколько предыдущих глав,— то все, сказанное относительно употребления и злоупотребления слов «наука» и «ученый», можно повторить и для слов «обучение» и «исследования». Чтобы покончить с этим призраком раз и навсегда, заявим совершенно определенно, что настоящая книга написана в духе афоризма «ты должен узнать истину, а эта истина должна сделать тебя свободным» и что в основе книги лежит вера в то, что накопление и распространение знаний является благороднейшей целью человека. Именно с этой точки зрения важно указать на ту степень коррупции, которой подверглись слова «наука» и «ученый». Быть может, эти термины когда-нибудь будут реабилитированы и восстановлены в своем былом величии, но эта перспектива маловероятна. Это может потребовать, например, введения строгих законов, запрещающих злоупотребление этими словами со стороны людей, пускающихся в сомнительные научные предприятия. Совершенно недостаточно того, что тот или иной человек сам называет себя ученым, а свою деятельность наукой. Необходимо двинуться дальше

* В смысле сведения сложного процесса к более простому.— *Прим. ред.*

этого и спросить его, что он знает, и уточнить, чем он занимается. Никакой закон не отнимет у астролога права относиться к своей астрологии как к науке. Но, с другой стороны, астрологией, как правило, занимаются мошенники и проходимцы при поддержке, быть может, сентиментальных антикваров. Аналогично те, кто настаивает на тождестве человека и машины, кто клянется, прибегая к научной терминологии, что через каких-нибудь десять лет вычислительные машины смогут доказывать важные математические теоремы, писать полноценную музыку, будут решать основные проблемы психологии и обыграют последнего чемпиона мира по шахматам, — это просто современные прорицатели.

«Всё прекрасное так же трудно, как и редко», — говорит Спиноза. То же самое можно сказать и о науке и о научных успехах в области вычислительных машин и обработки информации. Подлинные достижения так же трудны, как и редки. И еще реже они совпадают с результатами пророчества и преждевременных оповещений о том, что́ кто-то надеется получить, но еще не получил.

ЛИТЕРАТУРА

. J. Carter and G. Pollard, An Enquiry into the Nature of Certain Nineteenth Century Pamphlets, London, Constable and Co., Ltd., 1934, p. 94.

*Искусственный разум: прогресс и проблемы**

А. Л. Сэмюэль

Выражение «искусственный разум», которое напрашивается само собой, вошло во всеобщее употребление, не имея вполне определенного и всеми признанного значения. К сожалению, употребление этого термина несет известные антропоморфные представления, которые обычно вызывают эмоциональные реакции у читателей, мало знакомых с истинным положением дел. Для одних искусственный разум представляется научным заблуждением — Мифом о Думающих Машинах, для других искусственный разум — первые неумелые попытки человека создать машинные методы решения ряда обременительных задач, которые следовало бы с полным основанием передать машинам, но которые пока еще, видимо, требуют участия человеческого разума, и, наконец, для некоторых, слишком пугливых индивидуумов искусственный разум ассоциируется с надвигающейся опасностью господства Машины над человеком. Это расхождение мнений и чувств относительно предмета, как будто поддающегося научной оценке, свидетельствует об общем недостаточном знании, и настоящее обсуждение будет представлять собой попытку исправить создавшееся положение.

По существу, сейчас назревает революция в методах применения цифровых вычислительных машин для решения задач деловой жизни и промышленного производства. Эта революция берет свое начало в целом ряде научных исследований, которые, казалось бы, далеки друг от друга. Некоторые из них имеют на первый взгляд весьма далекие от экономики цели, например программирование вычислительных машин для игры в сравнение монет (match pennies) в крестики нулики, в шашки и шахматы, для сочинения стихов или музыки, решения задач по планиметрии из

---

* A. L. S a m u e l, Artifical intelligence: progress and problems, «Computers and automation», vol. XII, № 3, p. 28—35. Статья дается в переводе Г. Н. Поварова.— *Прим. ред.*

курса средней школы. Другие связаны с более практическими задачами — программированием вычислительных машин для чтения переданных от руки телеграфных сигналов, для распознавания письма и речи, для перевода с русского языка на английский. Целью третьих является создание обучающихся машин, механизация познавательных процессов и даже создание думающих машин.

Некоторое представление о размахе работы в этих областях можно получить из того факта, что одна недавно опубликованная библиография об искусственном разуме содержала 559 ссылок на статьи приблизительно 400 различных авторов и 26 дополнительных ссылок на материалы симпозиумов, сессий и прочие специальные сборники, посвященные проблеме искусственного разума, причем в этот перечень не вошли работы по вопросам механического перевода [1] (в более ранней библиографии 1959 года приводилось 645 статей на эту тему).

Многочисленная и все более растущая группа специалистов, главным образом в Соединенных Штатах (а также рассеянных и по всему цивилизованному миру), пробует составлять программы для существующих машин или проектировать и строить специальные машины с прямым назначением решать задачи, которые, будучи выполнены людьми или животными, считались бы свидетельством участия разума. Каковы бы ни были наши личные чувства по поводу моральных последствий этой формы автоматизации, мы не можем позволить себе игнорировать ее предстоящее революционизирующее воздействие на наше общество.

Среди «ультра»

Как всегда во всякой революции, и здесь есть свои «ультра» (lunatic fringe) — люди, верящие в чудо, те, кто одержим своим энтузиазмом к новому делу и выдвигает дикие проекты, угрожающие дискредитировать все предприятие. Среди работающих в области искусственного разума известно, быть может, слишком много таких людей. Норберт Винер, который, конечно, не принадлежит к их числу, способствовал тем не менее неточному мышлению, выдвигая как один из основных догматов кибернетики, что существует глубокая аналогия между функциями управления у людей и у машин, и подчеркивая ту мысль,

что машины могут обладать оригинальностью и что они составляют угрозу человечеству [2]. Напротив, Мортимер Таубе из Колумбийского университета относится критически ко всей этой области и даже обвиняет занятых в ней специалистов в «сочинительстве научных фантазий с целью возбудить публику и заработать себе легкие деньги или искусственную репутацию». Можем ли мы пройти через трясину мелких деталей и мешанину сталкивающихся претензий и так или иначе установить истинную природу искусственного разума и смысл надвигающейся революции?

Чтобы упростить нашу задачу, придется допустить некоторый произвол и отказаться от определенных применений цифровых вычислительных машин, таких, например, как расчет напряжений в крыле самолета или составление платежных ведомостей, даже если эти применения и уменьшают объем мыслительной работы, которая должна делаться человеком. В некотором смысле мыслительная работа в таких ситуациях выполняется людьми заранее, когда они записывают серию команд («программу» — на жаргоне вычислительной техники), которыми предопределяется последовательность отдельных шагов, необходимых в ходе вычислительного процесса.

Вычисления для управления станками

Поскольку даже эти банальные применения вычислительной машины могут показаться непосвященным несколько таинственными, мы, пожалуй, задержимся на минуту и подчеркнем то обстоятельство, что, в конце концов, цифровая вычислительная машина есть всего лишь неодушевленное собрание механических и электрических частей, действующее чисто механическим образом. Как отмечала свыше ста лет назад леди Лавлейс* в своем описании аналитической машины Чарлза Бэббеджа**,

---

* Леди Лавлейс, Ада Августа (Lovelace, 1815—1852), дочь поэта Байрона, была одной из немногих, поддерживавших идеи Бэббеджа по созданию аналитической машины.— *Прим. ред.*

** Чарлз Бэббедж (Babbage, 1792—1871) — английский математик и механик, один из первых изобретателей вычислительных машин с программным управлением. Главный проект его, проект «аналитической машины» (Analytical Engine), который ему удалось осуществить лишь частично, имел основные черты современных больших универсальных вычислительных машин.— *Прим. ред.*

от вычислительной машины нельзя получить ничего другого, кроме того, что в нее было вложено (исключая, конечно, редкие случаи неисправной работы), и вычислительная машина способна делать только то, что мы знаем, как объяснить ей делать.

Магия вычислительной машины не в том, что она может делать, а лишь в скорости и точности, с которыми ею выполняется последовательность очень простых вычислительных шагов, последовательность, которая, как мы уже говорили, задается заранее лицом, составляющим программу, и которая приобретает сложность и полезность от многообразных зависимостей, существующих между отдельными шагами. Эти банальные применения цифровых вычислительных машин характеризуются той особенностью, что ответы, которые должна давать вычислительная машина, получаются всегда из входных данных в результате применения точного набора правил, вполне известных и записанных заранее программистом. Эти правила могут быть строго логическими, как в научных проблемах, или могут грешить против логики, как наши законы о подоходном налоге, но они определены, и горе тому, кто их нарушит! В некотором смысле допустимо сказать, что ответы целиком содержатся во входных данных и что функция вычислительной машины сводится лишь к перегруппировке данных в более удобную форму путем применения определенных правил.

Программирование машины для таких вычислений даже в лучшем случае — трудная задача, но не столько из-за какой-либо внутренней сложности самой машины, сколько из-за необходимости предопределить каждый мелкий шаг процесса в самых мельчайших подробностях. Вычислительная машина, скажет вам любой программист,— это гигантский идиот и отнюдь не гигантские мозги! Когда вычислительную задачу дают человеку — ассистенту, ему говорят, что делать; когда же составляют программу для вычислительной машины, приходится говорить ей, как делать задачу. Такое различие между «что» и «как» далеко не тривиально. Оно отличает плохого служащего от хорошего и в гораздо большей степени определяет разницу между цифровой вычислительной машиной, как весьма эффективным, но крайне глупым помощником в вычислениях, и разумным человеческим существом.

Хотя цифровая вычислительная машина — это машина для работы с символами любого рода, ее история, как это удачно показал Аллен Ньюэлл из Технологического института имени Карнеги, начинается с настольной счетной машины и со счетно-аналитической машины. Однако, как уже раньше предлагал Бэббедж, для увеличения эффективности машины было внесено несколько новшеств. В частности, в машину были введены устройства трех специальных типов, в дополнение к арифметическому устройству.

Первое из них — память (правильнее, запасающее устройство*), где могут храниться команды и данные, пока они не понадобятся. Добавлены также входные и выходные устройства, позволяющие оператору вводить в машину команды и данные и получать запись желаемого выхода. Наконец, имеется устройство центрального управления, истолковывающее команды и выполняющее операции трех основных типов: перемещение информации с одного места на другое, выполнение простейших арифметических действий и передача управления команде, взятой по одному из двух или более заданных адресов, в зависимости от результата сравнения, проводимого этим устройством между знаками или относительными величинами заданных чисел. Эти команды имеют форму повелительных предложений, например: «Переместить X в адрес A; прибавить X к Y; пропустить одну команду, если X отрицательное», и т. д. Последовательность таких предложений, образующая программу, представляет собой, как мы уже говорили, описание, как решать задачу, а не указание, что решать. Коль скоро такое описание сделано, применять вычислительную машину к решению многих задач повседневной жизни становится шаблонным делом.

Как известно каждому, подобные банальные применения вычислительных машин имеют очень важное экономическое значение — настолько важное, что они уже вызывают революцию в наших методах проведения научных исследований, управления станками и организации бизнеса. Важно, однако, помнить, что эти перемены вызваны не каким-то мистическим процессом, доступным понима-

---

* Автор здесь пользуется терминами «memory» и «storage unit», считая второй более правильным. Мы переводим «memory» как «память» и «storage unit» как «запасающее устройство».— *Прим. перев.*

нию лишь немногих избранных, а прямым, хотя и кропотливым применением весьма элементарных процедур, к задачам, существенно рутинным по своей природе.

На границе

Наряду с этими простейшими задачами рутинного типа есть много умственных процессов, которые людям приходится выполнять, но которые не могут быть сведены — или до сих пор никем не были сведены — к простому набору правил. Возьмем процесс игры в шахматы — не простое соблюдение правил игры, а процесс ведения хорошей игры против разумного противника. Никаких известных процедур, гарантирующих выигрыш, здесь не существует, и тем не менее люди учатся играть в шахматы и некоторые становятся даже весьма искусными игроками.

Еще одним примером может служить проблема доказательства теорем в планиметрии. Забыв на минуту, что существует разрешающая процедура Тарского, которой ученики средней школы не знают, мы тем не менее способны развить в себе умение доказывать теоремы без помощи исчерпывающего процесса выписывания всех возможных цепочек логически выводимых предложений, который мог бы привести к доказательству*. Взамен этого принимается методика, где несколько более или менее произвольно выбранных процедур исследуются довольно неполным образом и каждая дает некоторый ключ для ориентировки, на правильном ли мы пути, пока наконец через серию догадок мы не придем к формулировке удовлетворительного доказательства. В обоих приведенных случаях мы можем иногда прийти к правильному или хотя бы к очень хорошему ответу за изумительно короткий отрезок времени, но вместе с тем отсутствует

---

* Здесь имеются в виду некоторые результаты математической логики. Разрешающая процедура, предложенная польским логиком А. Тарским (ныне проживающим в США), позволяет для каждого предложения (теоремы) элементарной планиметрии найти либо доказательство, либо опровержение, и притом за конечное число шагов. Процесс выписывания всех возможных цепочек логически выводимых предложений в поисках доказательства (или опровержения) данного предложения теоретически возможен, но, конечно, слишком громоздок, чтобы быть практически осуществимым; кроме того, он может продолжаться до бесконечности, не принося никакого результата.— *Прим. перев.*

уверенность в том, что мы вообще когда-либо получим решение, как и уверенность, что представляющееся нам решение наилучшее.

Такой метод решения задач получил название «эвристической» процедуры в отличие от применения «алгоритма». Последний термин используется в этой связи для обозначения полностью заданной процедуры решения, которая гарантирует получение ответа, если только мы затратим время на выполнение всех предписанных шагов. Заметим, что пробуемый, но несовершенный алгоритм сам по себе еще не является эвристической программой. Эвристическое решение задач, когда оно успешно, должно, конечно, расцениваться как более высокая умственная активность, чем решение задач посредством более или менее автоматической процедуры. Поэтому с нашей стороны, вероятно, будет оправдано подвести под категорию искусственного разума машинные методы и машинные программы, которые используют эвристические процедуры.

К решению этой проблемы искусственного разума существует два в корне различных подхода. Один подход (его мы и обсудим первым) состоит в анализе задач, которые, по-видимому, требуют участия человеческого разума, и затем в проектировании такой машины (или составлении соответствующей программы для существующей машины), которая, как мы ожидаем, решит эти задачи. Специфические механизмы решения подобных задач в человеческом мозгу нас не интересуют; мы анализируем задачу, а не устройство, которое ее решает.

Обратимся к избитой уже аналогии. Когда человек впервые пытался летать, он изучал птиц, и первые неудачные летательные аппараты были механическими птицами. В основном прогресс был достигнут лишь тогда, когда человек перестал копировать полет птиц и начал изучать аэродинамику. Современные реактивные самолеты должны преодолевать в полете те же самые аэродинамические проблемы, что и птицы, однако механизмы, используемые птицами при решении проблемы полета, совсем другие.

Наблюдения над птицами

Другой подход — изучение птиц, а не аэродинамики — имеет свои достоинства, и в области искусственного разума, где столь многое еще неизвестно, на этом пути

были получены весьма интересные результаты. Я имею в виду общее направление исследований, основанное на работах канадского психолога Дональда Хебба в университете имени Мак Гилла, развитое Натаниелем Рочестером из Международной корпорации деловых машин (фирма IBM), Фарли и Кларком из Массачусетского технологического института и ставшее недавно популярным благодаря Франку Розенблатту из Корнелльского университета, создавшему термин «перцептрон».

Мысли этого направления примерно таковы. Мозг человека, подобно мозгу животных, состоит из большого числа клеток определенного типа, называемых нейронами. Эти клетки обладают довольно необычными свойствами; они реагируют по принципу «все или ничего» (на профессиональном жаргоне «возбуждаются») и передают импульс другим нейронам через синаптические связи. Каждый нейрон соединен со многими другими, и для возбуждения нейронов обычно требуется поступление нескольких входных сигналов.

Насколько можно определить, имеется известный порядок в общей картине взаимосвязей между нейронами, но также, по-видимому, налицо и некоторая степень случайности в деталях соединений. Обучение, вероятно, состоит в изменениях в силе и, может быть, даже в числе этих синаптических взаимосвязей.

В настоящее время возможно придумать механические, химические и электрические устройства, грубо имитирующие поведение отдельных нейронов, и мы можем соединить эти устройства друг с другом случайным образом, чтобы имитировать синаптические взаимосвязи, существующие в мозгу, и наконец можем предусмотреть автоматическое усиление или ослабление этих связей с помощью тренировочной программы.

Хотя степень разума, достигнутая до сих пор, при этом фактически еще очень низка, такие устройства обладают некоторыми весьма интересными свойствами. Во-первых, их можно использовать при решении задач, для которых мы не располагаем полной математической формулировкой. Во-вторых, они в достаточной мере универсальны. Например, устройство, построенное для распознавания характерных отметок на бумаге, может быть обучено распознавать английские буквы. Обучение могло бы состоять в подаче устройству серии входных сигналов (в нашем

случае букв алфавита), в усилении тех частных внутренних взаимосвязей, которые побуждают машину давать правильный ответ, и в ослаблении связей, ведущих к неправильным ответам. Такой процесс вполне аналогичен методу поощрений и наказаний, применяемому при дрессировке животных, при этом весьма важно то обстоятельство, что данное устройство можно было бы столь же хорошо обучить распознавать китайские иероглифы или, скажем, определять характерные географические подробности на аэрофотоснимке.

За недостатком места позволим себе только упомянуть еще одно интересное свойство рассматриваемых устройств — предполагаемую экономию элементов, по крайней мере в отношении записывания информации,— экономию, которая, по-видимому, должна создаваться тем, что информация запасается в соединениях между элементами, а не в самих элементах. Мы вернемся к этой интересной теме после того, как рассмотрим «аэродинамику» нашей проблемы.

Назад к аэродинамике

Марвин Минский из Массачусетского технологического института, обсуждая проблему искусственного разума [3], решил разбить тему обсуждения на пять отдельных областей: поиск, распознавание образов, обучение, планирование и индукцию. Хотя это деление до некоторой степени произвольно, оно расчленяет проблему так, что, как нам кажется, позволяет схватить существенные черты и, кроме того, показывает, что здесь нет никакой магии. Вместо этого нам предстоит обсудить цепочку простых шагов, которая может быть механизирована. Вряд ли можно дать лучшее резюме проблемы, чем у Минского.

«Вычислительная машина может делать в известном смысле лишь то, что ей задано делать. Но если даже мы не знаем точно, как решать задачу, мы можем запрограммировать машину, чтобы она обследовала некоторое большое пространство попыток решения [т. е. испытывала много, много решений, одно за другим]. К сожалению, когда мы составляем прямую процедуру для такого поиска, то находим результирующий процесс чрезвычайно

неэффективным. [Решений оказывается слишком много.] С помощью техники распознавания образов эффективность можно значительно повысить, позволяя машине применять свои методы только к тем попыткам, для которых они целесообразны. А при обучении эффективность машины повышается еще больше, так как поиск направляется прежним опытом. Анализируя текущую ситуацию и применяя то, что мы называем методами планирования, машина может добиться действительно фундаментального улучшения, выражающегося в замене первоначально заданного поиска гораздо более сжатым, более целесообразным исследованием».

Далее Минский упоминает о некоторых, как он называет, более глобальных понятиях, относящихся к индукции, и здесь мы можем пойти дальше своими собственными путями.

Перейдя к реальной обстановке, рассмотрим задачу, которая является более или менее тривиальной,— задачу программирования вычислительной машины для игры в шашки. Я выбрал шашки, а не шахматы отчасти по личным причинам (так как сам составлял такую программу [4]), отчасти из-за очевидной простоты игры (такой пример проливает яркий свет на проблему поиска). Чтобы заставить вычислительную машину играть в шашки, мы прежде всего должны изобразить положение шашек на доске по способу, который вычислительная машина могла бы запомнить. Затем последствия каждого из имеющихся ходов должны анализироваться заглядыванием в будущее, как это в общем делал бы человек, рассматривая каждый начальный ход по очереди, затем — все возможные ответные ходы противника, затем для каждого из них — все контрответы и т. д.

Средний человек способен продолжить этот процесс заглядывания в будущее только на два-три хода вперед, но кто-нибудь мог бы утверждать, что поскольку вычислительные машины работают чрезвычайно быстро, то все же, вероятно, удалось бы проследить все возможные ходы вплоть до самого конца игры и таким путем совершенно однозначно установить относительную ценность возможных первых ходов.

К сожалению, вычислительные машины работают не так уж быстро. Даже если мысленно перенестись в будущее, к наибыстрейшей возможной машине, подчи-

ненной лишь таким ограничениям, как размеры вселенной, молекулярная природа материи и конечная скорость света, то и такой вычислительной машине потребовались бы многие столетия, а может быть, и более длительный срок, чем даже возраст вселенной, чтобы сделать свой первый ход.

Горовосхождение

Человек, решая задачу очередного хода, перестает в надлежащий момент заглядывать в будущее и оценивает результирующие положения на доске, руководствуясь некоторыми ближайшими целями: имел ли он возможность взять одну из фигур противника без потери своей фигуры? Имел ли он возможность провести простую шашку в дамки? Или даже имел ли он возможность разыграть дебют, ведущий к дамочным полям? Такой анализ, естественно, не может быть исчерпывающим, и подобные промежуточные цели не совсем надежны в качестве указаний, что мы следуем в правильном направлении. Эта общая процедура получила название горовосхождения*, и присущие ей недостатки становятся очевидными, если попытаться применить ее для того, чтобы достичь вершины Эвереста, идя все время вверх и отправляясь, скажем, из Гарден-Сити на Лонг-Айленде.

Налицо две трудности: во-первых, существование местных пиков (если холм Вест-Хилл в 380 футов над уровнем моря заслуживает такого наименования) и, во-вторых, существование областей, имеющих лишь некоторые небольшие неровности (например, поверхность океана), где мы начинаем бесцельно петлять, если шаги не будут велики, и в то же время можем совсем пропустить нужный нам пик, если шаги не будут малы. Не приходится говорить, что для того, чтобы справиться с такими проблемами, было разработано много различных методов; однако в той или иной степени такие проблемы всегда могут существовать, так же как, видимо, они всегда досаждают человеку.

Окончив частную горовосходительную экскурсию или частное заглядывание в будущее при игре в шашки, мы

---

* Термином «горовосхождение» мы передаем английский термин hill climbing; в русской литературе это понятие обозначается также словами «подъем по градиенту».— *Прим. перев.*

должны затем определить высоту места или оценить положение на доске, чтобы увидеть, стали ли мы хоть немного ближе к своей цели. При действительном подъеме на горы существуют такие приборы, как высотомеры, которыми можно измерить точную высоту над уровнем моря, но в аналогичном случае с шашками у нас нет никакого простого способа для оценки качества шашечной позиции. Конечно, существует возможность, что точно такая же позиция уже встречалась нам в одной из предыдущих партий, но это крайне невероятно, исключая, пожалуй, дебюты и эндшпили, и потому вместо прямой оценки мы вынуждены ограничиться типизацией шашечной ситуации, то есть классифицировать ее как принадлежащую к некоторому общему классу шашечных ситуаций. Задача не лишена сходства с задачей распознания в пятнах или мазках типографской краски на печатной странице образа, который мы воспринимаем как печатную букву, например как букву А. Это проблема распознания образов.

Распознание образов включает два шага. Первый состоит в создании некоторого числа понятий, например понятия круглости, используемого при распознании букв О, понятий прямизны и направленности линий и т. п. Второй шаг состоит в приписывании весов этим различным свойствам, когда они используются при распознании или классификации неизвестных букв. Делались попытки механизировать оба эти шага, но до сих пор достигнут лишь незначительный прогресс в отношении первого из них — образования понятий, и большинство исследователей довольствовались подбором выработанных у людей понятий и разработкой машинных процедур для приписывания весов этим понятиям.

Обучение машины

Здесь-то мы и встречаемся с идеей обучения машины. Предположим, мы изобрели какой-то автоматический способ проверки эффективности любого заданного распределения весов в отношении фактической точности распознания и подготовили механизм, который меняет распределение весов так, чтобы получалась максимальная точность. Не требуется входить в подробности такой

процедуры, чтобы увидеть, что она может выполняться полностью автоматически и что машина, соответственным образом запрограммированная, будет «учиться» на собственном опыте.

Одним из подтверждающих фактов является то, что программа игры в шашки (о которой я упоминал выше) имеет как раз такой характер, и весьма легко показать, что программа эта «учится» на своем же собственном опыте и что со временем она становится все более совершенным шашистом. Это свойство имеет то забавное следствие, что программа вскоре делается способной побить человека, который составлял ее, не потому, что она пользуется какой-либо информацией или методикой, неизвестной или недоступной человеку, а лишь благодаря своей непогрешимой памяти, фантастической аккуратности и чудесной скорости, которые позволяют ей проделать за несколько секунд детальный, но не озаренный воображением анализ, занявший бы у человека годы.

78 Мы углубились во все эти подробности не для того, чтобы сделать из наших читателей профессиональных программистов, но чтобы показать, насколько прозаическим является в действительности весь предмет и насколько далекими кажутся наши разумные машины от разумного человеческого существа. Достаточно напомнить, что мастер-шашист все еще может побить лучшую шашечную программу, несмотря на свою жалкую по сравнению с машинными стандартами память и разницу более чем в миллион раз в относительных скоростях вычислений. Программы обучения предстоит еще только применить к чему-нибудь более сложному, чем шашки, а реальные задачи, к которым нам хотелось бы обратиться, имеют сложность на много порядков выше.

Парадокс

Пожалуй, нами сказано достаточно в отношении отрицательных сторон этого вопроса. В обучении машин происходит прогресс, и нам когда-нибудь станет понятно, почему человек может превзойти машину, а в результате такого понимания мы сумеем спроектировать лучшие машины или же запрограммировать существующие машины так, что они смогут превзойти человека в большей части форм умственной деятельности. В самом деле,

можно подозревать, что и наши нынешние машины были бы способны сделать это, если бы только мы были достаточно находчивы в составлении программ такого рода. Препятствующие этому ограничения лежат не в машине, а в самом человеке.

Но тогда возникает парадокс. Чтобы сделать машины, которые казались бы более находчивыми, чем человек, человек сам должен быть более находчивым, чем машина. По-видимому, для того чтобы сообщить машине искусство быть разумной, требуется разум (или хотя бы понимание) более высокого порядка, чем для того, чтобы дублировать разум, тот разум, который машина должна лишь имитировать.

Когда мы наконец достигнем этой степени понимания, необходимой для составления программы, подражающей человеку в большей части форм его умственной деятельности, тогда мы почувствуем надобность в составлении еще более обобщенной программы, которая бы заставила машину составлять программы для самой себя или для другой машины. Это в свою очередь потребует еще большего понимания со стороны человека. Такому процессу нет конца, но, очевидно, человек, как инициатор, всегда будет впереди.

Наш общий вывод, следовательно, гласит, что нам нечего опасаться машины, по крайней мере в отношении того, что существует какая-либо опасность появления машины, которая оказалась бы разумнее человека. Разум машины предписан человеком, и для составления такого предписания необходима более высокая степень мышления, чем для его выполнения.

Подумав поглубже

В нашем рассуждении, однако, содержится ошибка. Мы предполагали, что человек не будет способен построить разумную машину до тех пор, пока он не сможет полностью понять внутренней работы такого устройства. Тем не менее на всем протяжении истории человек открывал многие свойства природы, которых он не понимал, и, несмотря на отсутствие понимания, умел использовать их как для добра, так и для зла. Мы должны, следовательно, считаться с возможностью того, что чудовище, подобное Франкенштейну, будет создано человеком в виде

разумной машины более или менее случайно и задолго до того, как человек выработает подробное знание, необходимое для контроля создания его рук.

Это возвращает нас к другому методу решения проблемы искусственного разума, о котором мы говорили раньше, — методу изучения птиц, а не изучения аэродинамики, ибо кажется разумным предположить, что из названных двух альтернатив этот метод имеет бо́льшую вероятность привести к открытию без понимания. Напомним, что процедура состоит в имитации мозга с помощью случайно скоммутированной сети нейроноподобных устройств в надежде, что такой ансамбль будет обладать разумом. Мы могли бы далее утверждать, что так как подробности соединений между элементами будут неизвестны, то должна существовать какая-то степень неопределенности в нашем знании способностей ансамбля и в нашей способности предсказания его поведения и в этой мере наше устройство может оказаться более рузумным, чем человек, его создатель.

Это, однако, крайне невероятно. Во-первых, встает проблема относительной величины мозга и машины. Соединять нейроноподобные устройства становится все труднее и труднее, когда их делается намного больше, чем, скажем, 10^6 или 10^7 элементов, и в этих условиях отдельные устройства по необходимости должны быть очень простыми. Между тем мозг человека содержит, возможно, 10^{10} нейронов, и отдельные нейроны имеют много отростков, которые соединяют их очень сложным образом с другими нейронами. В среднем, вероятно, имеется сотня таких соединений на нейрон. Перед лицом этой сложности наши слабые попытки имитации в гораздо большей мере напоминают нервную систему плоского червя, чем повторенный мозг человека. Такое положение, правда, сохранится не всегда, и с помощью будущих методов изготовления вычислительных машин мы, возможно, и будем когда-нибудь способны создавать устройства, приближающиеся по сложности к мозгу.

Один шанс из миллиона миллионов

Второе обстоятельство связано с нашим недостаточным знанием детальной организации межнейронных соединений в мозгу. Мозг наверняка не скоммутирован случайно.

хотя можно выставить доводы и в пользу обратного утверждения. Например, когда сложность измеряется величинами порядка 10^{20}, трудно себе представить, каким образом все эти детали могут храниться в зародышевой плазме. Имеются также данные о большой избыточности мозга, так как известны случаи, когда относительно большие участки мозга были повреждены или даже вовсе удалены без сколько-нибудь серьезного долговременного ухудшения умственных способностей.

С другой стороны, наблюдается некоторая глобальная организация, присущая мозгу всех животных. Определенная организация требуется, очевидно, и для того, чтобы наделить новорожденного ребенка столькими рефлексами и инстинктивными схемами поведения, сколько необходимо для его выживания. О наличии упорядоченной организации свидетельствуют также многие мелкие психические черты, которые мы наследуем. Наконец, хотя гениальные люди в среднем как будто имеют более крупный мозг, чем их менее счастливые собратья, это наверняка не общее правило, и, по-видимому, главное различие между умственно одаренными и умственно отсталыми скрывается скорее в детальном строении мозга, а не просто в его объеме.

Конечно, поскольку об организации мозга мы, по существу, ничего не знаем, предположение о случайности — столь же хороший отправной пункт, как и любое другое предположение, однако наши шансы сконструировать устройство, сходное с мозгом человека, составляют тогда что-то порядка 1 к 10^{12}. Только благодаря более глубокому пониманию основных механизмов затронутых явлений мы сможем достичь лучших результатов и будем постепенно совершенствовать нашу способность контроля*.

Мы можем, следовательно, повторить свое предыдущее заключение, что нам нечего опасаться со стороны машины в отношении ее возможного господства над человеком. Но это отнюдь не означает, что человек не сможет использовать вычислительную машину во вред человечеству.

Цифровая вычислительная машина, которая имеется сегодня, и разумная машина, которая появится завтра,—

---

* Автор не учитывает возможности самоорганизации и самосовершенствования имитирующих мозг устройств в порядке ускоренного прохождения эволюции, что требует специального рассмотрения.— *Прим. перев.*

орудия, точно так же, как пишущая машинка, экскаватор и термоядерная бомба, а большая часть человеческих орудий, если не все они, может быть использована как святыми, так и грешниками. Цифровой вычислительной машине может недоставать огромной разрушительной силы бомбы, но зато бомба своей эффективностью во многом обязана вычислениям, сделанным машинами, и в некотором смысле вычислительная машина должна будет разделить вину, если человек преуспеет в уничтожении самого себя. Однако последнюю ответственность должен будет нести сам человек, и попытки приписать вину неодушевленному собранию механических и электрических частей, которые человек соединил или заставил соединиться, представляют собой лишь жалкую попытку свалить свои грехи на кого-нибудь другого.

*Музыкальные кресла**

Однако полностью избавиться от угрозы появления разумной машины не так легко, как могло бы показаться после этих замечаний. Хотя машина, возможно, и никогда не будет разумнее человека, но люди очень отличаются друг от друга по своим умственным способностям, и машина может и, несомненно, должна превзойти каких-то людей. Итак, угроза заключается скорее в технологической безработице, чем в господстве машины над человеком. В некотором смысле это лишь продолжение процесса, который происходит все время с тех пор, как человек сделал свое первое изобретение.

Констатация этого факта не решает проблемы, но влияние разумной машины на занятость рабочей силы должно бы рассматриваться точно так же, как влияние любой другой формы автоматизации, и не должно выделяться особо. Правда, искусственный разум — это нечто большее, чем еще одно новое средство сбережения труда,

* «Музыкальные кресла»— игра, состоящая в том, что участники маршируют под музыку вокруг стульев, число которых на единицу меньше числа участников. Как только музыка прекращается, каждый спешит сесть на стул, а оставшийся без места выбывает из игры. После каждого тура один стул удаляется, пока наконец не останется лишь двое участников и единственный стул. Тогда последний тур определит победителя. В данном случае заголовок намекает на угрозу безработицы (сокращение рабочих мест) из-за внедрения разумных машин.— *Прим. перев.*

ибо он усиливает мозг человека, а не его мускулы. Действительно, искусственный разум вполне способен оказаться тем новым фактором в нашей экономике, который позволит человеку решить всю проблему технологической незанятости, включая и долю, если она вообще будет иметь место, которую вызвало его собственное внедрение*.

Программирование вычислительных машин для ведения игр, сочинения стихов и решения задач по планиметрии из курса средней школы представляет собой лишь стадию в развитии понимания тех методов, которые необходимо будет применить для машинной имитации разумного поведения. Мы еще находимся на стадии игры, но по мере того, как мы будем идти вперед в понимании, кажется разумным предположить, что новые методы будут применяться к ситуациям реальной жизни все чаще и чаще и что усилия, посвященные проблеме игр и другим детским проблемам, будут уменьшаться. Быть может, мы еще не достигли этого поворотного пункта и должны еще узнать многое. Тем не менее вряд ли можно сомневаться, что уже недалеко время, когда основная масса более тривиальных умственных работ, отнимающих сейчас у человека столько времени, будет делаться машиной. Искусственный разум — не миф и не угроза человеку.

ЛИТЕРАТУРА

1. Marvin Minsky, A Selected Bibliography to the Literature on Artificial Intelligence, «IRE Transactions on Human Factors in Electronics», 2, March, 1961, p. 39—55.
2. Norbert Wiener, Cybernetics; or Control and Communication in the Animal and the Machine, New York, John Wiley & Sons, 1948 (Н. Винер, Кибернетика, изд. Советское радио, 1958); Some Moral and Technical Consequences of Automation, «Science,» 131, May, 1960, p. 1355. См. также Arthur L. Samuel, Some Moral and Technical Consequences of Automation — A Refutation, «Science», 132, September, 1960, p. 741.
3. Marvin Minsky, Steps Towards Artificial Intelligence, Proc. I. R. E.», 49, January, 1961, p. 8—30 (М. Минский, На пути к искусственному мышлению, «Труды института радиоинженеров», № 1, 1961, стр. 13—36).
4. Arthur L. Samuel, Some Studies in Machine Learning Using the Game of Checkers, IBM J. of Research and Development, 3, No. 3, July, 1959, p. 211.

---

* Для советского читателя ясно, что эта проблема не технологическая, а социальная и связана с природой капитализма.— *Прим. перев.*

*Кибернетика сегодня и ее будущий вклад в технические науки**

У. Росс Эшби

Надо взять на себя большую смелость, чтобы ответить сегодня на вопросы, в каком состоянии находится кибернетика и как она будет развиваться, так как при этом затрагиваются не только вопросы будущего науки, но также скромная работа большого числа квалифицированных работников. Я готов взять на себя такую ответственность и заявить, что позиция сегодня достаточно ясна, для того чтобы обосновать решительные утверждения по этим вопросам.

По-моему, кибернетика сейчас настолько выкристаллизовалась, что можно охватить ее целиком и увидеть пределы ее возможностей. Что это за пределы — об этом я попытаюсь сказать немного позже. Они кажутся мне совершенно ясными, и я уверен, что кибернетика в течение многих лет развития будет двигаться в рамках определенных идей. Между тем эти идеи широко известны в течение уже нескольких лет и могли бы навлечь на себя критику и быть опровергнуты. Ничего этого не происходит. Поэтому у меня сложилось мнение, что эти точки зрения продержатся, вероятно, достаточное время, несмотря на то, что некоторые из них требуют решительного изменения в наших способах мышления.

Давайте обратимся к нашей главной теме. Я полностью согласен с теми, кто начинает с утверждения, что целью мира техники является управление. Это, во-первых, вопросы практического управления (часто называемые общим термином «исследование операций»); во-вторых, более абстрактная и общая теория управления и, наконец, это «путеводный свет», или «блуждающий огонек», кибернетики: попытка разработать процессы управления, подобные процессам, происходящим в мозгу человека.

---

* W. Ross Ashby, Cybernetics Today and its Future Contribution to the Engineering-Sciences, New York, 1961. Статья дается в переводе Ю. Г. Дадаева.— *Прим. ред.*

О первых двух вопросах я не скажу ничего: по существу, они являются простыми. А вот кибернетику я попытаюсь охарактеризовать и показать, каково ее значение. Оправдает ли она наши надежды в будущем или окажется простой иллюзией, основанной на ошибочной аналогии?

Современное состояние

Для того чтобы сделать ясным обсуждаемый предмет, необходимо прежде всего, чтобы мы поняли, что человеческий мозг действительно существует и что он действительно работает. Это необходимо как тем, кто решает, какими должны быть цели кибернетических исследований, так и тем, кто будет их выполнять. Еще в древней Греции эта область была наводнена мифами и различными религиозными предрассудками. И только буквально в наши дни научно изучается человеческий мозг и его истинное поведение. Раньше психология занималась «изящным» обсуждением законов логики, основанным на предположении, что все обсуждающееся является таким видом мышления, которым можно гордиться. Действительным процессам, которые происходят в наших головах, со всеми их ошибками, путаницей, колебаниями, отклонениями и прочими подобными недостатками, не оказывалось внимания. И в результате греки и следующие их учению психологи развили теорию о том, как мозг *должен был бы* работать. Эти идеи явились для нас общим наследием, и с ним нелегко справиться. Между тем слишком много работающих в области кибернетики используют идеи о работе мозга, которые действительно сродни теориям средневековых схоластов. Они используют мощную современную технику для того, чтобы ответить на вопросы, которые никогда не ставились, потому что нельзя ждать от мозга подобных действий. Мы бы только посмеялись, если бы современный химик предложил для разложения земли, воздуха, огня и воды использовать урановый котел. Сегодняшние кибернетические исследования еще в большей степени открыты для критики подобного рода. Чтобы не быть голословным, разрешите мне привести несколько примеров — мифов о мозге, существующих еще и доныне.

Одна распространенная идея, например, касается способности мозга предугадывать будущее. Идея, конечно, внешне правдоподобна, но по существу ошибочна. Винер

кратко изложил истину, когда сказал, что предсказание будущего — это действие, основанное на прошлом. Если прошлое повторяется в будущем и если будущее поддерживает такую непрерывность, то мозг будет достигать успеха. Но этот процесс полностью основан на прошлых событиях и на степени проявления последовательной регулярности реального мира.

Пусть природа создаст что-то действительно новое, и мозг становится беспомощным. Что произошло, например, с теми, кто работал с рентгеновыми лучами в 1905 году? Они получили страшные ожоги. Мозг не может действовать эффективно, столкнувшись с новым явлением. Он должен ждать, пока пройдет достаточное время и новое останется в прошлом. Всякая мудрость — это мудрость после происшедшего события. И все, что мы ожидаем от механического мозга, должно быть основано на этом факте.

Еще один миф. Мы задаем вопрос: «Как мозг выполняет обобщения?»—и обычно получаем ответ: «Он не выполняет их». Рассмотрим, например, опыты Пэйперта. Он носил призматические очки и, таким образом, видел окружающий его мир, в том числе и части своего тела, зеркально отображенными. Вытягивая руки, он следил, когда кто-нибудь коснется его руки, и пытался делать движение рукой, мешающее прикосновению. После нескольких ошибочных движений он стал искусным и развил новую связь «глаз — рука». Потом он научился проделывать то же самое со ступнями ног, и ему показалось, что он полностью обучился перемене сторон — правой и левой. Тогда он решил проделать тот же опыт с коленями. Было обнаружено, что у доктора Пэйперта, который отнюдь не страдал от недостатка ума и сообразительности, обобщения реакции на перемену сторон не произошло, и надо было специально научиться проделывать этот опыт и относительно колен. Мозг не обладает универсальной способностью «обобщения».

В качестве последнего примера рассмотрим миф о том, что определенные люди обладают особым свойством «гениальности», которое делает их в некотором смысле совершеннее других людей, несмотря на общность всех других свойств. Миф основан скорее всего на плохой статистике. Многие работники берутся за решение задачи, используя все накопленные ранее идеи о том, как решить ее; кто-то добивается успеха, и тогда *мы* выделяем его

и предполагаем, что он должен обладать чем-то, что одинаково хорошо для всех времен и для всех задач. В XVII веке, например, многие математики стали в тупик перед исчислением бесконечно малых. Ньютон подошел к задаче со свойственной ему склонностью мыслить обо всем как о чем-то непрерывном, переходящем от одного состояния к другому. Удивительно ли, что он вместе с Лейбницем благоприятствовал открытию? А смог бы он добиться такого же успеха как физик в начале этого столетия, когда была необходимость в ком-то, кто мог легко представить себе дискретные, скачкообразные изменения в атомах? Многие представления о «гении» возникают из-за элементарной, но грубой ошибки, что мы сами, дескать, крепки задним умом, а вот можем назвать человека, который действительно сделал открытие первым. Борясь с подобными заблуждениями, нужно сказать, что гении, упорные и плодовитые,— это в большинстве случаев люди, которые овладели своей специальностью и посвящают ей бо́льшую часть суток в отличие от «средних» людей. Если Моцарт видит сломанную ветку, то этот образ тотчас же вызывает в нем новую музыкальную фразу; крик птицы дает ему идеи новых гармоний; из всего извлекается «музыкальная польза», и за один час, казалось бы, «ничегонеделания» он мог достичь в музыке большего, чем многие из нас за всю жизнь. И удивительно ли, что он мог писать симфонии на заказ?

Эти несколько примеров дают некоторые указания на степень пересмотра нашего сегодняшнего представления о мозге, необходимого для того, чтобы иметь реальную картину мозга и успешно создать искусственный мозг. Однако эта оценка заблуждений является только негативной. Более важным является позитивное понимание полезности мозга во всех его проявлениях, которые подчиняются одному основному правилу. Поэтому разрешите мне от средневековых религиозных предрассудков перейти к краткому изложению понятий о мозге 1961 года.

Что такое мозг?

Имеется, конечно, много аспектов рассмотрения живого мозга. Для хирурга это — ком студня, который при сотрясении вибрирует определенным образом. Однако здесь будет идти речь об особом аспекте, когда живой

мозг демонстрирует то, что мы называем умением. Он всегда в курсе дела; он достигает результатов; он попадает в самые трудные ситуации и, находя успешные решения, демонстрирует свою мощь. Эту способность мозга мы все пытаемся познать, особенно в мире техники, когда беремся за создание механического мозга. Что такое мозг, когда он достигнет своей замечательной победы над природой?

Замечание, которое я сейчас сделаю, личного порядка, но оно может оказаться полезным. Когда я начинал работать в 1928 году, я предполагал, что к мозгу требуется отмычка; стоит найти ее, и мы откроем путь к «чистому разуму», готовому к немедленному использованию в неограниченных количествах, подобно нефтяному фонтану. Сегодня нас не проведешь: мы знаем, что нет такой отмычки, и не потому, что мы отчаялись найти ее, а потому, что мы теперь твердо знаем, что́ представляет собой мозг. Эти новые данные могут быть описаны довольно легко, несмотря на необходимость соблюдать осторожность, в особенности по отношению к новым знаниям, применяемым к старым идеям и не являющимся простыми модификациями последних.

Первый факт заключается в том, что обладание любым умением или способностью развивается под действием или естественного отбора, или процессов обучения и образования, воздействующих на индивидуум в детстве и в последующие годы. После миллиарда (или около того) лет формирования под действием окружающей среды мозг в настоящее время приспособлен к окружающей среде с поразительной точностью. Но он приспособлен именно к этой окружающей среде. Такая приспособленность не более универсальна, чем приспособляемость системы управления процессом переработки нефти или системы управления пуском ракет.

Далее, живой мозг приспосабливался для того, чтобы выжить; и это предусматривало постоянное достижение определенных целей: необходимо было добывать пищу, утолять жажду, избегать ран, перегрева или охлаждения и т. д. Разум проявляется в этих или иных, дополнительных, целях. Таким образом, первый принцип систем, подобных мозгу, или мыслящих систем, живых, или искусственных сводится к следующему:

Система может быть подобной мозгу, или разумной, только по отношению к определенной задаче.

Другими словами, конструирование систем, подобных мозгу, не может быть начато до тех пор, пока нет ответа на вопрос: «Что вы хотите?»

Таким образом, сегодняшняя точка зрения на моделирование мозга, исходящая из того, что это прежде всего управление, является пробным камнем. Это урок двадцати лет развития кибернетики. Придерживаться этой точки зрения — значит находиться на современных научных позициях. Если забыть о такой точке зрения, то можно скатиться к чему-то мифическому, сверхъестественному, одним словом, к совершенно неправдоподобному.

Я говорю «неправдоподобное», потому что это подчеркивает строгость основ наших знаний. Результаты Шеннона относительно коррекции шумов впервые были применены в телефонных и радиоканалах. Несмотря на такое применение, они могут быть использованы и для других целей, внешне совершенно не схожих с вопросами телефонной связи, так как эти теоремы по своему содержанию полностью абстрактны. Действительно, они имеют огромное применение для всех процессов управления: коррекция шумов здесь превращается в коррекцию любого отклонения от того, что нужно получить. Эти теоремы устанавливают точный верхний предел того, что может быть достигнуто с помощью какого-то процесса с учетом обычной связи между причиной и следствием. Другими словами, если мозг работает при обычных связях между причиной и следствием (и без серьезных сомнений в наличии этих связей со стороны работника), то достижения его как целенаправленного устройства имеют точный предел. Подобно тому, как работа мускулов человека ограничена сверху законом сохранения энергии, так и деятельность мозга ограничена теоремами Шеннона* и законом необходимого разнообразия**.

В свете этого нового понимания у нас есть достаточное основание рассмотреть вопрос: «Можем ли мы создать сверхмозг?» Но, во-первых, мы должны быть уверены, что

* Имеется в виду главным образом десятая теорема Шеннона, связывающая возможность коррекции ошибок с пропускной способностью канала. См. К. Шеннон, Работы по теории информации и кибернетике, Издательство иностранной литературы, 1963, стр. 278.—*Прим. ред.*

** О законе необходимого разнообразия см. У. Росс Эшби, Введение в кибернетику, Издательство иностранной литературы, 1959, стр. 293—302, 346—350.— *Прим. ред.*

правильно ставим вопрос. Неправильно (как это имело место в 1928 году) думать, что мозг имеет какую-то отмычку, при помощи которой стоит только открыть его, и получим лампу Аладина: взмах руки — и будешь иметь все, что хочешь. Эта точка зрения перестает существовать под ударами теорем Шеннона и закона необходимого разнообразия. Для того чтобы получить супермозг, не надо искать отмычки.

Что же тогда остается? Нам кажется, что вопрос, сможем ли мы создать мозг, лучший, чем мозг человека, является незначительным, не более интересным, чем те вопросы, которые возникли сто или более лет назад, когда инженеры сделали первый электромотор мощностью $^1/_7$ *л. с.*, получив таким образом мотор, превышающий мощность человеческих мускулов. Очевидно, когда они изобрели мотор, то вопрос превышения мощности мускулов человека имел скорее сентиментальную окраску, чем технический интерес. Подобная ситуация существует и сегодня при создании системы со способностью мышления, превышающей способности человека, который создает эту систему.

Сегодня создание системы, подобной мозгу, с определенными интеллектуальными способностями в принципе так же доступно, как создание паровой машины с заданной мощностью. В действительности можно набросать схему такой работы. Во-первых, должна быть определена цель (которая может быть комплексной и временной); далее, необходимо произвести идентификацию факторов, имеющих тенденцию помешать достижению этой цели; и затем шаг за шагом переходить к самому созданию регулятора, который будет так обрабатывать информацию, что, несмотря на имеющиеся помехи, цель будет достигнута.

Если этот процесс может показаться кое-кому слишком простым и обычным и если кое-кто думает, что мозг человека является более гибким, то мне хотелось бы просить привести действительное доказательство, на котором основано такое недоверие. Как я указывал ранее, идеи о том, что определенные люди являются «гениальными», исчезают при внимательном рассмотрении фактов. Могу только сказать, что у меня нет ни одного факта, который позволил бы предположить, что человеческий мозг достигает результатов иным путем, чем тот, о котором только что говорилось.

Подобно тому, как энергетика зиждется на законе сохранения энергии — каждый эрг работы, который машина может затратить, должен быть получен, — так и любой регулирующий и управляющий прибор основан на законе, по которому достигаемая степень соответствующего отбора ограничена количеством принятой и обрабатываемой информации. Таким образом, мы можем создать искусственный мозг сколь угодно большой и совершенный, обеспеченный достаточным количеством обрабатываемой информации. Кое-кто скажет, что мы нашли лампу Аладина и знаем теперь, что с ее помощью можно получить все что угодно. Но это одна сторона медали. Все что мы получаем, должно быть оплачено!

Разрешите мне еще раз вкратце обрисовать сегодняшнее положение. В течение веков мозг, как и сердце, наделялся множеством мифических свойств. В конечном счете было обнаружено, что сердце — это насос, так же обстоит дело и с мозгом, который является регулятором и, следовательно, устройством для обработки информации. Он имеет ограничения, накладываемые теоремами Шеннона и законом необходимого разнообразия. Огромная сила человеческого мозга заключается в его бесконечной приспособляемости к окружающим условиям действительности. Однако нет причин предполагать, что любой другой механизм обработки данных с подобной способностью к эволюционной работе должен быть хуже. Также нет причин предполагать, что неживые механизмы не способны превзойти мозг в совершенстве своих действий. Уровень действий, которого может достигнуть механический мозг, прямо зависит от количества проектной и экспериментальной работы, затраченной на его создание.

Будущее кибернетики

Как я уже говорил, в кибернетике сейчас выкристаллизовалось ясное и устойчивое понимание основных принципов создания механизмов, подобных мозгу. Хорошо или плохо, эти основные принципы будут служить нам в качестве научной основы на протяжении многих лет, до тех пор, пока, как это всегда случается, не произойдет замена их чем-нибудь лучшим. Но хотя кибернетика, как я указал, достигла времени своей «кристаллизации», это никоим образом не означает, что она перестанет развиваться. И я хочу рассмотреть ее будущее влияние

на развитие техники, постановку задач специального образования и проведение исследовательской работы.

Начнем с будущего кибернетики в технике, обсудим, будет ли кибернетика активной и животворной силой или останется всего лишь набором академических теорий.

Я глубоко верю, что в свое время вклад кибернетики в технику будет убедительным и это произойдет ранее чем через 10 или более лет, как это ожидается. Что мне кажется выдающимся — это открытие *ограничений*, наложенных на искусственный мозг теоремами Шеннона и законом необходимого разнообразия.

Ситуация в этом вопросе подобна той, которая была сто лет назад, когда энергетики почти ежегодно изобретали новые машины и каждый инженер с богатым воображением надеялся, что вот-вот появится наконец и вечный двигатель. Потом обнаружилась неприятная неожиданность. Доказательств тому, что все машины существенно ограничены, что энергия сохраняется, что каждый эрг работы, который затрачивается, прежде должен быть получен, становилось все больше. В то время многие инженеры чувствовали глубокое разочарование. Тем не менее *принятие* этих ограничений было первым шагом к новому и в общем более реальному пониманию принципов машиностроения. На протяжении всего времени инженеры, которые приняли эти ограничения, перегоняют тех, кто остается верным надежде изобрести вечный двигатель, и именно потому, что идеи первых в большей степени согласуются с действительными возможностями. Поэтому я, несомненно, чувствую, что те, кто принимает современные идеи о мозге, об его ограничениях и ограничениях всех механизмов, подобных мозгу, будут в конечном счете впереди тех, кто цепляется за старую идею о том, что живой мозг обладает какой-то таинственной силой, позволяющей сделать все в мгновение ока.

Чем же будет характерен новый метод? Коль скоро принимается, что все действия, подобные действиям мозга, сводятся к действиям соответствующего отбора, достигаемого с помощью обработки информации, мы можем сосредоточить внимание на важном вопросе — на эффективности системы. Мы работаем теперь, не занимаясь поисками магического действия, которое выполняет все в мгновение ока, а совершенствуя те процессы, с помощью которых достигается цель с разумной степенью эффективности.

Это изменение, внешне незначительное, на самом деле имеет большие последствия. Действительно, мы еще далеко не сведущи в вопросах эффективности; в то время как энергетики смеются сегодня над системой с коэффициентом полезного действия, равным 10%, мы оказываемся порой совершенно невежественными, используя информационные процессы, обладающие коэффициентом полезного действия, немного меньшим 1%. Чтобы это не казалось чрезмерным преувеличением, разрешите мне сослаться на один конкретный пример.

После пятидесяти лет обсуждения, в котором участвовали сотни людей (включая и меня), только сейчас стало ясно, а именно благодаря Дону Кэмпбеллу с Северо-Запада, что если мы знаем, какова цель, то мы можем, используя эти знания, сократить время ее поиска. Но насколько? Если мы обладаем знаниями о цели в начальной стадии и используем эту информацию, то число действий сокращается до корня квадратного от первоначального числа. Таким образом, если первоначальное число было миллион операций, то предложенный метод сокращает это число до тысячи. Это означает, что, если мы имеем информацию о цели, но не используем ее, коэффициент полезного действия нашей работы в этом примере составляет 0,1%! Такая неэффективность встречается сегодня сплошь и рядом в наших методах работы с процессами, подобными процессам в мозге.

Мы должны согласиться, что сегодня наше интуитивное чувство относительно количества информации почти не совершенствуется. Будем надеяться, что молодое поколение, более привыкшее к работе с дискретными величинами, чем с непрерывными функциями, будет постоянно изменять подобное положение. Ясно, что нельзя достигнуть основательного увеличения эффективности процессов, подобных процессам в мозге, до тех пор пока на всех стадиях искусственных процессов не будет достигнута эффективность, по крайней мере одного порядка величины с той, какую обнаруживает мозг.

Вкратце, каково же будущее кибернетики в технике?

Прежде всего надо отбросить мысли о неуловимых свойствах мозга, о магических силах мышления и прийти к действительному пониманию работы мозга. Это означает практически, что надо начинать с определенной частной задачи, а затем приступать к последовательному созданию

механизма управления, сосредоточив внимание на максимизации эффективности.

Это, может быть, звучит слишком общо, но это так. Когда инженер начинает работать над созданием газовой турбины или ракеты с ядерным зарядом, он не ждет чудес, а приступает к последовательному выполнению программы этих работ, постепенно продвигаясь от, возможно, слишком низкого, первоначального значения эффективности к наивысшему. Таким же должен быть и кибернетический метод в технике для создания систем, обладающих мощью живого мозга.

Чему надо учить?

Из моих разговоров о том, как кибернетика будет использоваться в технике, видно, чему мы должны учить наших студентов; даже минутное размышление заставляет нас признать, что сегодня наша система обучения далеко не удовлетворительна. Обучение специальным предметам, таким, как теория сервомеханизмов или теория телефонной связи, отвечает современным требованиям и является высококвалифицированным. Но по-иному обстоит дело с теорией «разумных» устройств. Мы до некоторой степени страдаем сегодня от исторической случайности. Всегда ожидали, что нейрофизиологи в конечном счете сделают открытие в области работы мозга. Но пришла вторая мировая война, и активность заметно упала. Оживление вновь началось в 1947 году, и тут нейрофизиологи столкнулись лицом к лицу с непонятными машинами, которые выполняли процессы, подобные процессам в мозгу, и с непонятной математикой — теорией информации, которая была нужна для описания этих процессов. За некоторым исключением, нейрофизиологи забросили сей предмет, забаррикадировавшись микроскопами. Если вы думаете, что это слишком сильно сказано, то я позволю себе обратить ваше внимание на новый справочник по физиологии, который должен рассматриваться как лучшее, что могло быть выпущено сейчас. Нейрофизиологии посвящено три солидных тома, заполненных самыми научными сведениями, тем не менее они совершенно ничем не могут помочь умному молодому человеку, который хочет только узнать: «Что такое мозг?» Ведь это факт, что в вопросах физиологии высшей нервной деятельности большинство нейрофизиологов отстают на пять-десять лет.

Каким же образом учить кибернетике?

Ответить на этот вопрос нелегко; отчасти это административный вопрос, так как кибернетика неравномерно разделилась на более чем полдюжины дисциплин. Специальные дисциплины сравнительно легко распределить по факультетам, труднее обстоит дело с основами наук, которые сейчас разбросаны по разным факультетам. При решении этого вопроса желательными компонентами являются простота, здравый смысл и полное исключение технических деталей до тех пор, пока необходимость в них не станет очевидной. Такое обучение должно начинаться, я думаю, с изучения эволюции, естественного отбора и гомеостаза; при этом термин «мозг» не будет иметь четкого определения в его функциональном смысле. Потом следует переходить к общей теории сервомеханизмов и регуляторов (не делая упор на линейные случаи) и постепенно — к важному понятию количества информации при усложнении системы. Количество информации может оцениваться комбинаторным способом или с помощью метода Мак-Хилла и Гарнера, которые более удобны, чем специальные методы Шеннона. После всего этого студент способен судить о состоянии предмета и решить, в каком направлении ему специализироваться. Кстати, мне кажется, что в большинстве университетов курс обучения следовало бы переделать, распределив его между несколькими преподавателями, чтобы каждый имел дело с определенной дисциплиной; таким образом, они охватили бы все необходимые предметы. Подобный курс обучения дал бы студенту основу знаний для восприятия физиологии высшей нервной деятельности (если он специализируется по психологии или физиологии) и для того, чтобы понять, как строить мозгоподобные разумные машины (если он специализируется по кибернетике или вообще по техническим дисциплинам). Я рекомендую такой путь обучения с достаточной уверенностью, так как я применял его уже в Урбане, и, собственно говоря, он найден был на практике.

Дальнейшие исследования в области кибернетики

В заключение мне бы хотелось сказать кое-что о будущих исследованиях в области кибернетики. Здесь я дол-

жен говорить в первую очередь, очевидно, что-то критическое.

Как мне кажется, некоторые из современных работ характеризуются скорее сложностью техники, чем значительностью их целей. Сложнейшая современная техника подчас используется для решения задач, которые являются мелкими и даже средневековыми по своей концепции. Что бы мы сказали, если бы медики-исследователи захотели вдруг применить масс-спектрограф для исследования того, какой из четырех темпераментов главным образом присущ индивидууму. Это смешно. Однако некоторые исследования, проводимые сегодня, можно отнести только к подобному типу исследований. Такая ситуация, конечно, изживет сама себя, но в настоящее время имеется настоятельная необходимость в тщательном отборе методов исследований в работах по кибернетике, особенно при решении ее основных задач. Я должен быть осторожным, чтобы не обидеть настоящих работников, но не могу удержаться от совета, чтобы все, кто имеет отношение к таким работам, будь то руководитель или исполнитель, настойчиво спрашивали себя: как рассматривается мозг в этих исследованиях? С позиций 1961 или 1941 года? Даже за этот короткий срок произошла полная революция в наших взглядах на высшие функции мозга, и эта революция должна быть отражена в исследованиях сегодняшнего дня.

Что можно сказать, я повторяюсь, когда исследования обращены к некоторой предполагаемой таинственной силе мозга? В 1961 году такой подход оправдать нельзя, и он является таким же вредным, как, скажем, ссылка на какие-то таинственные возможности электричества с целью изменения цен на бирже. Сегодня должно действовать правило: каждое исследование, которое имеет отношение к свойствам, подобным свойствам мозга, должно говорить, какие это свойства и почему именно они необходимы. В конечном итоге обоснование должно быть в некотором смысле практическим и дано в общепринятых выражениях. Обращение сегодня к *мистическому* толкованию мозга отличает шарлатана или, может быть, невежду.

В проблемах исследований нет недостатка. Например, как я отмечал ранее, мы слишком мало знаем об эффективности различных регулирующих процессов, особенно

примитивных, но в то же время основных, встречающихся в биологических системах и действующих методом проб и ошибок.

Далее, мы практически ничего не знаем об общих свойствах систем со многими состояниями равновесия, систем, если можно так выразиться, с распределенной памятью. К последнему классу систем следует отнести системы с порогом срабатывания; это кажется почти невероятным, но, несмотря на то, что за 50 лет было выяснено, что многие ячейки мозга работают на принципе порогового срабатывания, мы слишком мало знаем об общей динамике систем подобного типа.

Другая насущная необходимость — это необходимость в формальной математической теории упрощения. В будущем прогресс в отношении мозга-гиганта будет сильно зависеть от нашей способности упрощать. В настоящее время упрощения выполняются интуитивно с помощью естественной сообразительности. Нам необходимо создать теорию «упрощений». Работы математиков по гомоморфизму ведутся в непосредственной близости к этому, но их работу необходимо переориентировать в соответствии с нуждами теории информации и технических наук.

Суммирую: если по некоторым вопросам я говорил чересчур резко, то не думайте, что я полностью осуждаю их. В 1928 году, когда я впервые занялся мозгом и его высшими функциями, предмет сей был окружен непроницаемой тайной. Сегодня эта тайна исчезла, и кибернетика находится в трудном переходном состоянии. От некоего научно-фантастического отправного пункта она движется к сплоченности и единогласию. Сегодня мы готовы совершенствовать и применять новые знания. Ничто не стоит на нашем пути, разве только наше типично человеческое свойство цепляться за свои старые идеи даже после того, как они оказываются устарелыми. Будущее — в исследовании и управлении механизмами, подобными мозгу,— в руках тех, кто способен решительно пересмотреть свои старые взгляды на мозг, и будущего поколения. Если оно должным образом будет обучено, то сможет взглянуть на мозг реально и прийти к доказательству того, что наука о механизмах, подобных мозгу, в высшей степени понятна, целесообразна и полезна.

*О возможностях кибернетических систем**

Поль Армер

Птица — это аппарат, действующий согласно математическим законам. И воспроизвести этот аппарат со всеми его движениями — в пределах возможностей человека.

Леонардо да Винчи

В этой статье я попытаюсь проанализировать некоторые взгляды и споры, возникшие в связи с такими вопросами, как «Может ли машина мыслить?» или «Может ли машина действовать разумно?» Я делаю это с единственной целью — содействовать улучшению климата, окружающего исследования в области машинного, или искусственного, мозга. Я не собираюсь убеждать тех, кто отрицательно отвечает на эти вопросы, в том, что они неправы, а правы те, кто отвечает на них положительно. Просто я попробую показать, что большинство разногласий — это вопрос семантики. Я попытаюсь опровергнуть некоторые доводы «негативистов»**, считающих исследования в области искусственного мозга ненужными, и надеюсь убедить их в том, что им следовало бы быть более терпимыми к таким исследованиям. Тех, кто отвечает утвердительно на эти вопросы, мне ни в чем не придется убеждать. Однако если они разделяют мое мнение в отношении важности исследований в данной области, то, мне кажется, мы можем плодотворно провести время в обсуждениях наших взглядов, потому что существующие сегодня взгляды «негативистов» тормозят эти исследования [1].

---

* Сообщение сделано на Симпозиуме по бионике (WADD Technical Report 60, 600, 1961).

Статья дается с некоторыми сокращениями.— *Прим. ред.*

** Здесь и далее под термином «негативист» автор понимает сторонника отрицательного решения вопроса о возможности реализации искусственного разума; решающих этот вопрос положительно автор называет «позитивистами». — *Прим. ред.*

История вопроса

Прежде чем изучать сущность этих споров и взглядов, рассмотрим кое-что из истории обсуждаемого вопроса, так как идея машин, наделенных разумом, существует уже давно.

Сэмюэль Батлер (1835—1902) описал настоящую гражданскую войну между «машинистами» и «антимашинистами» [2]. (Между прочим, тогда победили «антимашинисты».) Батлер заявлял: «Нет гарантии против развития в конечном счете искусственного сознания на базе существующих в наше время машин, обладающих некоторыми признаками сознания», и предполагал, что может настать время, когда «человек будет для этих машин тем же, чем для нас являются лошадь и собака». Этот вопрос, очевидно, обсуждался и во времена Бэббеджа, потому что леди Лавлейс комментировала его в своих записках о работах Бэббеджа [3], правда отрицательно. Идея разумных машин вновь всплыла в конце 40-х годов нашего века, когда с появлением первых больших электронных цифровых вычислительных машин (ЭЦВМ) осуществились мечты Бэббеджа. Когда же популярная пресса стала применять к этим машинам термин «гигантский мозг», конструкторы и эксплуатационники этих машин, в том числе и я, немедленно почти все до единого встали на защиту человеческого разума. Мы поспешили заявить, что вычислительные машины не «мыслят»; они лишь с большой скоростью выполняют арифметические операции. Ажиотаж вокруг этого вопроса затих (но не совсем) в начале и середине 50-х годов, но вспыхнул с невиданной ранее силой в последние несколько лет. Он вторгся и на страницы журнала «Сайенс» [1, 4, 5].

Доводы «негативистов»

Изучение возражений, выдвинутых «негативистами», показывает, что многие из этих возражений вовсе не являются доводами, а всего лишь декларациями. Некоторые из этих заявлений совершенно неосновательны, например «Давайте договоримся раз и навсегда, что машины не могут мыслить!» или «ЭЦВМ — это совсем не гигантский мозг, как пытаются вас уверить некоторые воскресные приложения и авторы научно-фантастических рас-

сказов. Это чрезвычайно быстрый и исключительно точный «дурак»» [6].

Многие доводы оказались ошибочными, типа «Машины никогда не смогут стать разумными, потому что они обладают (или не обладают) тем-то и тем-то свойством». Источники ошибки здесь могут быть разными. Наличие или отсутствие определенных свойств может быть случайным, или же из наличия или отсутствия тех или иных свойств логически совершенно не следует вывод «негативиста». Либо «негативист» заблуждается, полагая неизменными какие-либо свойства. Позвольте мне привести некоторые примеры.

«Манчестерская машина при решении шахматных задач, по-видимому, руководствовалась этим методом, а именно просматривала все возможные последовательности всех возможных ходов. Это, между прочим, раскрывает все достоинства и все недостатки конструкции машины. Она может просмотреть гораздо больше возможностей за данный отрезок времени, чем человек, но она должна просмотреть все возможности. Человек же воспринимает доску как единое целое и интуитивно отвергает множество возможностей. Машина не может делать ни того, ни другого» [7].

Эти утверждения относительно поведения машины совершенно несправедливы. Верно, что некоторые первые методы конструирования машин, играющих в шахматы, предлагали просмотр всех последовательностей ходов ограниченной длины [8], но это не единственный способ решения данной проблемы. Программы игры в шахматы Ньювела, Шоу и Симона [9[не предусматривают исследования всех возможностей без разбора. А те возможности, которые рассматриваются, машина исследует с различной степенью подробности. Она отбрасывает плохие ходы, отбирает ходы, кажущиеся хорошими, и глубоко исследует их, чтобы установить, действительно ли они являются хорошими. Более ранняя программа, разработанная той же группой авторов для доказательства теорем логики [10], не перебирала всех возможных доказательств, в противном случае для решения такой задачи на современных ЭЦВМ потребовалась бы вечность. Вернее, она искала в лабиринте возможных доказательств те, которые выглядели подходящими, и исследовала их. Она использовала знание того, какие направления уже просмотрены.

Пример ошибочного допущения того, что свойства машины неизменны, имеется в статье Дж. Х. Тролла.

«Человеческая память — это регистрирующая система, имеющая значительно большую емкость, чем запоминающее устройство самой большой из имеющихся кибернетических машин. Механический мозг, который имел бы столько же ламп и реле, сколько в человеческом мозге нервных клеток (около 10 млрд.), не поместился бы в Эмпайр Стейт Билдинг* и потребовал бы полной мощности Ниагарского водопада для своего питания и всей реки Ниагары для своего охлаждения. Больше того, такая ЭЦВМ не проработает и доли секунды, как выйдут из строя несколько тысяч ламп, которые потребуется заменить» [11].

Выводы Тролла привязаны к электронной вакуумной лампе (статья была написана в 1954 году) и поэтому были скомпрометированы уже с появлением транзистора, который занимает меньше места, более экономичен и значительно более надежен, чем вакуумная лампа. Вместе с тем, правда, оказалось, что Тролл назвал, несомненно, слишком заниженные цифры для числа нервных клеток. Однако современная техника тонких пленок открывает возможность изготовления деталей для кибернетической машины, невидимых в оптический микроскоп [12]. Имеются основания ожидать, что окажется возможным с помощью этой техники разместить в 1 кубическом футе** пространства такое же количество логических элементов, какое содержится в мозгу человека. Питание потребуется незначительное.

Другие возражения «негативистов» несут в себе заблуждения семантического характера. В дальнейшем я затрону их в неявной форме.

«Позитивисты» и понятие мышления

Разумеется, существует много «позитивистов». Это прежде всего Джон фон Нейман, Норберт Винер, А. М.Тьюринг.

А. М. Тьюринг был одним из первых, кто глубоко рассмотрел проблемы машинного мышления (имеется в виду

* Самый большой небоскреб Нью-Йорка (105 этажей). — *Прим. ред.*

** 1 кубический фут $\approx$ 0,0298 $м^3$.— *Прим. ред.*

его научная статья, опубликованная в 1950 году) [13]. Хотя он и рассматривает основной вопрос — может ли у машины проявиться способность мыслить, но признает трудность точного определения слов «машина» и «мышление». Чтобы обойти эту проблему, он вместо нее рассматривает игру, в которой играющий, связанный с человеком и с машиной посредством телетайпа, должен определить, с кем (или с чем) он связан в данный момент. Чтение этой статьи — настоящее удовольствие для тех, кто интересуется разумными машинами. Так же как и Тьюринг, я избежал определения слова «мыслить». И хотя я не буду пытаться делать этого в явной форме, мне все же хотелось бы развить одну идею для описания мышления, которая, я надеюсь, поможет преодолеть некоторые семантические трудности, связанные с изучением вопроса: «Может ли машина мыслить?» Идея следующая: мышление — это континуум, n-мерный континуум. Эта идея определенно не новая, она существует с тех пор, как человек впервые сравнил свои умственные способности со способностями другого человека. Она неявно присутствует во всех аргументах «позитивистов» в отношении разумных машин. Психологи давным-давно ввели «частное интеллектуальности»* в качестве меры в этом континууме, и их идея «факторов» указывает на n-мерность континуума разумности. Использование одномерного «частного интеллектуальности», очевидно, является сверхупрощением того, что есть на самом деле. Хотя эта идея и не нова, я все же не знаю примера, где о ней явно говорилось бы в контексте о разумных машинах. И это меня всегда удивляло.

Позвольте мне провести аналогию с континуумом транспортной способности. В отношении скорости перевозки людей из Нью-Йорка в Лос-Анжелес современный реактивный самолет «заткнет за пояс» все остальные существующие виды транспорта. Но перевозить бумагу для газет из Британской Колумбии в Калифорнию по стои-

* В оригинале — intelligence quotient (частное интеллектуальности) — отношение числа очков, которое индивидуум набирает в данном умственном тесте, к числу очков, которое средний индивидуум его возраста набрал бы в том же самом тесте; частное интеллектуальности колеблется в пределах от 0,4 до 1,5; для удобства его умножают на 100. Введено Уэкслером: см. D. Wechsler Measurement of Adult Intelligence, Baltimore, Williams and Wilkins, 3-rd Ed., 1944.— *Прим. ред.*

мости выгоднее пароходом, чем самолетом. Самолет в свою очередь не может перевозить людей от одного дома к другому. Кадиллак может быть самым удобным средством для перевозки людей на небольшие расстояния по хорошим дорогам, но он едва ли заменит джип в условиях бездорожья: достоинства джипа — проходимость и выносливость. В этом измерении континуума транспортной способности человек вытесняет джип, потому что человек пройдет там, где джип не сможет пройти, так же как и джип сможет пройти там, где кадиллак не пройдет.

Подобным образом мы можем провести сравнение между людьми и машинами в континууме мышления. Если вы возражаете против использования слова «мышление», можно заменить его на «обработку информации» или что-нибудь в этом роде. Но я полагаю, что вы должны допустить существование некоторого континуума поведения, в котором люди и машины существуют. И в отношении некоторых измерений машина действует лучше человека, например она может выполнять арифметические операции значительно быстрее и надежнее. В свете такого подхода к рассмотрению поведения людей и машин довод, будто машина не может играть в шахматы, так как «она могла бы только оперировать с фигурами стандартного размера и типа и не смогла бы распознать множество фигур различного изготовления, тогда как человек распознает их и ходит весьма просто» [5], подобен тому, если сказать, что самолет братьев Райт не мог бы летать, так как он не мог бы пролететь без посадки из Лос-Анжелеса в Нью-Йорк и не мог сесть на дерево, как птица. Почему критерий разумности должен требовать, чтобы машина достигла в точности той же точки в континууме, что и человек? Является ли критерием способности летать достижение той же точки в континууме передвижения, которой достигла птица?

Многие «негативисты» заявляют, что в качестве единственного свидетельства разумности машины они примут достижение в нашем континууме, редко удающееся человеку. Например, они преуменьшают значение усилий в направлении машинного сочинения музыки, потому что то, что достигнуто в этой области сейчас, нельзя сравнить с творением Моцарта или Шопена. А многие ли композиторы смогут написать что-либо подобное? Крайний аргумент такого рода был приведен недавно на одном

собрании в Англии, на котором выступавший заявил, что он не согласится с тем, что машины могут мыслить до тех пор, пока машина не докажет знаменитого предложения Ферма, известного как последняя теорема Ферма. По логике выступавшего можно также заключить, что до сих пор человек не был способен мыслить, так как эта теорема остается недоказанной.

Некоторые «негативисты» хотят включить в свое определение разумного поведения требование того, чтобы оно могло приписываться исключительно живым организмам. Если следовать этому определению, то я должен признать, что машины не могут быть разумными. Мне кажется, что этот довод является семантическим, потому что мы знаем примеры поведения машины, которое можно сравнить с поведением человека. И утверждать, что работы по моделированию такого поведения человека с помощью машины не нужно,— все равно, что сделать вывод о бесполезности работ по созданию искусственного сердца на основании того факта, что искусственный орган — это не живой орган.

На мой взгляд, если так понимать мышление, как я понимаю, машина может мыслить и разумно вести себя. Если ребенок складывает два и два, мы обычно считаем его поведение разумным, особенно если это наш собственный ребенок. Тем из вас, кто возразит мне, что, мол, машина не знает и не понимает, что она делает, я приведу пример с ребенком, который как попугай повторяет детские стишки. Подобно этому, когда инженер, основываясь на образовании и опыте, проектирует электродвигатель, мы обычно характеризуем его поведение как разумное. А если машина делает ту же работу — разве это не разумно? Конечно, способности машины очень ограниченны по сравнению со способностями инженера, который может выполнять не только расчет двигателей по жесткой методике, но и многие другие вещи. Аналогично джипы не могут взобраться на Эверест, а самолеты ДС-8 не могут сесть на дерево.

Сводный список задач, которые в настоящее время могут решаться кибернетической машиной, но которые, как мы обычно считаем, требуют сообразительности, когда их решает человек, включает такие задачи, как:

1) доказательство теорем логики и планиметрии [10, 14];
2) игра в шашки и шахматы [15, 9];

3) управление сборочным конвейером [16];
4) сочинение музыки [17];
5) расчет двигателей [18];
6) дешифрирование кода Морзе [19].

Кибернетические машины постепенно овладевают теми способностями, которые в прошлом приписывались исключительно человеку. Время и место не позволяют мне приводить дальнейшие свидетельства машинного разума, но все это можно найти как в статьях, цитированных выше, так и в других работах [20, 21, 22]. Такие факты, выдвинутые «позитивистами», несомненно, служат основой для многих аргументов.

Для того чтобы доказать, что машины *сегодня не* проявляют разумности, необходимо лишь определить нижнюю границу в нашем континууме, о котором много говорилось ранее, и сказать, что, если поведение объекта выше этой границы, оно разумно, если ниже, не разумно. Я подчеркиваю, что многие «негативисты», придерживающиеся такого начала, каждый раз передвигают эту границу, так что она все время остается выше того, на что сегодня способны машины. Они (негативисты) могут делать это несознательно; так, например, Марвин Минский пишет:

«Разумность — это химера. Вы считаете данное действие разумным до тех пор, пока его не поймете. После объяснения вам все становится ясно. Если вы все же его не поймете, оно, по-видимому, разумно» [23].

Все мы заметили, что у других искусство выполнять недоступную нам работу представляется нам очень трудным. Когда же мы сами овладеваем этим искусством, оно становится для нас значительно менее внушительным. Было бы полезно на будущее иметь под рукой несколько вех, чтобы определить меру химеры Минского. Тьюринг поставил одну такую веху, но этого мало. Для этой цели нам нужна четко сформулированная задача, представляющая в настоящий момент исключительно объект деятельности человека (и поэтому она должна быть, несомненно, из области мышления), но такая, чтобы мы могли надеяться на ее решение в будущем с помощью машины.

Даже если вы допустите, что и человек, и машина могут действовать разумно, все еще остается вопрос: «Существует ли в нашем континууме граница, которую машина никогда не перейдет?» Лично я не готов ответить ни да, ни нет. Те из вас, кто верит в бессмертие человека

и в существование верховного существа, не сомневаются в существовании верхней границы для машины. Однако я настоятельно подчеркиваю, что в отношении вопроса важности исследований в области дальнейшего развития машин в нашем континууме не важно, существует верхняя граница или нет. Если же вы доказываете, что такая граница все же существует, тогда, конечно, интересен вопрос, где она проходит. Но я считаю, что мы имеем весьма мало данных о положении этой границы. Именно по вопросу о расположении этой верхней границы я и обращаюсь к здравому смыслу тех, кто ратует за то, что машины никогда не станут разумными. Ведь очень легко слишком занизить эту границу. Мне вспоминается, как кое-кто возражал против дальнейшего развития автомобиля, основывая свое возражение на том, что человек физиологически неспособен вынести скоростей, превышающих 15 миль в час.

Тролл, упоминавшийся ранее как автор примера с Эмпайр Стейт Билдинг, вмещающим 10 млрд. радиоламп, хотя и разделял мнение, что машины могут мыслить, в своей статье провел верхнюю границу, положение которой основано на ошибочном допущении, что свойства машины (габариты, энергия питания, энергия охлаждения, надежность) являются неизменными. Тролл, очевидно, не представлял себе времени, когда мы сможем построить машину с тем же числом элементов, что имеется в мозгу человека, небольшого размера, очень надежную и с низкими энергетическими требованиями.

Конечно, большого количества элементов, высокой надежности и малого расхода энергии — всего этого еще недостаточно. Теперь приобретает первостепенное значение организация элементов. Но разве можно считать невозможным, что мы узнаем, как организовать эти элементы (или, в предельном случае, узнать, как заставить эти элементы самих организоваться) так, чтобы получить машину, гораздо более развитую в континууме разумности, чем сегодняшние машины (хотя и не достигшую предельного совершенства)? Разве такая машина не будет представлять огромную ценность? Кстати, сегодняшние машины превосходят человеческий мозг в некоторых отношениях (например, в скорости выполнения арифметических операций) и, следовательно, представляют сегодня для человека значительную ценность.

Позвольте мне проиллюстрировать свою точку зрения следующим примером. Нам всем приходилось когда-либо пытаться вспомнить имя человека, которого мы когда-то встречали. В частности, д-р Уиллис Вэр и я пытались вспомнить имя одного человека. Мы перечисляли друг другу его приметы, где он работает, чем занимается и т. д., но имени его никак не могли вспомнить. Спустя некоторое время я повернулся к д-ру Вэру и сказал: «Его имя начинается на букву «З»». Тогда д-р Вэр щелкнул пальцами и произнес: «Так вот, его зовут Фрайзел!» Ценность машин, которые обладали бы подобной способностью обработки информации, была бы несомненной при решении проблем отыскания или восстановления нужной информации, для решения некоторых военных задач, интеллектуальных задач и т. п. А разве это не мыслимо, что мы в один прекрасный день поймем логику процессов, протекших в голове д-ра Вэра, и поэтому сможем реализовать их с помощью машины?

Мы, очевидно, не будем иметь такой машины до тех пор, пока кто-нибудь не поверит, что ее можно создать, не попытается построить ее. Чтобы попытаться построить акую машину, вам совсем не обязательно уверовать то, что верхняя граница, о которой много говорилось ыше, не существует, а всего лишь нужно согласиться тем, что она расположена значительно выше положения, которое занимают сегодняшние машины. Я не могу строго доказать, что эта граница не существует, но я и не встречал до сих пор доказательств, что это не так.

Здесь я хотел бы перефразировать цитату да Винчи, которой начиналась эта статья, и несколько расширить ее в интересах самой же истории. Таким образом, да Винчи мог бы сказать:

«Когда человек поймет естественные законы, которым подчиняется полет птицы, он сможет построить летательный аппарат».

Человек потратил огромное количество времени и сил, стараясь построить летательную машину, машущую крыльями, подобно птице, но важно то, что это был путь к открытию закона аэродинамической подъемной силы крыла, закона, который в итоге позволил человеку построить летательную машину. Эта аналогия довольно глубокая. Полет — это тоже континуум. Некоторые считали, что есть граница (скорость звука), за которой полет

невозможен, и т. п. Я надеюсь, что более важные стороны этой аналогии не нуждаются в пояснениях.

Отношение к исследованиям

Я считаю, что совсем не обязательно, прежде чем приступать к какому-либо исследованию, доказывать предварительно, что цель этого исследования достижима. Однако мне кажется, многие «негативисты» считают это необходимым в отношении исследований в области искусственного разума. Как я уже сказал, хотя я и не могу строго доказать, что граница в нашем континууме, выше которой машина не может подняться, не существует, в то же время нет сведений о ее точном местонахождении. А поскольку ее положение неизвестно и даже нет ничего определенного относительно ее существования вообще, то исследования, направленные на дальнейшее совершенствование машин в нашем континууме, выглядят вполне законными.

«Соперничество» между человеком и машиной

Я думаю, что позиция многих «негативистов» обусловлена влиянием сильного психологического фактора. Я уверен, что это главный фактор, заставивший меня стать «негативистом» десять лет назад. Согласиться с тем, что машина может обладать разумом,— это значит признать, что у человека есть соперник в области, в которой он доселе был единовластным хозяином. Тролл в своей статье, упоминавшейся выше [11], довольно подробно рассматривает это соперничество. Разрешите мне процитировать отрывок из письма, полученного нашим институтом, чтобы пояснить эту точку зрения:

«...семантика может с большим энтузиазмом многое сделать для поддержания исследований в этой области [в области искусственного разума]. Что касается лично меня, то термины «разумная машина» и «мыслящая машина» беспокоят меня и даже кажутся мне немного угрожающими. Я — человек и поэтому «разумен», а эти бесчеловечные аппараты собираются состязаться со мной и могут даже победить меня. С другой стороны, если бы те же самые «черные ящики» назывались «разрешителями

проблем» или даже «реализаторами приемлемого решения проблем», то они казались бы мне более дружелюбными, будь они способны помочь мне самым эффективным способом делать то, что мне нужно, но лучше все же, чтобы я был хозяином. Эти мысли целиком субъективны и чувственны...»

Позвольте мне также привести пример того, как «негативист» пытается объяснить психологию «позитивистов».

«Использование этих выражений [мыслящие машины и искусственный мозг], которые предполагают, что эти машины в некотором смысле обнаруживают признаки разума, символизирует существование широко распространенного, хотя и непризнанного желания верить в то, что машина может быть чем-то большим, чем просто машина» [24].

Ответ на вопрос, почему некоторые «негативисты» тщательно прощупывают свой путь, следует увязать с тем, что можно было бы назвать «грехами «позитивистов»». Преувеличения при популяризации достижений, особенно со стороны рекламных ведомств и фирм, производящих кибернетические машины, вызвали такую сильную реакцию в научных кругах, что многие впали в противоположную крайность.

Разумные машины и современные вычислительные устройства

Общее мнение таково, что сегодняшние вычислительные устройства являются исключительно арифметическими приборами. Хотя это и правда, что первоначально эти машины были построены для выполнения арифметических операций, они способны решать и другие, нечисловые задачи. Сущность вычислительных устройств состоит в манипуляции символами, и первое использование в них числовых символов явилось лишь исторической случайностью. Этот неправильный взгляд на вычислительные машины как на исключительно цифровые приборы привел к тому, что многие не могут себе вообразить такие устройства способными разумно себя вести, так как потребовалось бы (по их мнению) свести этот процесс к арифметическому. *Только прошу не истолковывать моих замечаний в том смысле, будто я приравниваю человеческий мозг цифровым вычислительным машинам в том виде, какими*

мы их знаем сегодня. В то время как я считаю, что сегодняшние цифровые вычислительные машины могут демонстрировать разумное поведение, я совсем не утверждаю, что разумные машины 1970-х годов должны обязательно напоминать сегодняшние или функционально, или физически. В частности, в моем желании видеть машины более развитыми в континууме разумности меньше всего участвуют интересы быстродействия, а, очевидно, гораздо более важными являются вопросы организации. Аналогично я совсем не придерживаюсь исключительно дискретного (цифрового) подхода: вполне возможно, что наиболее удачным окажется сочетание аналоговой и дискретной аппаратуры. Вместе с тем я не собираюсь отрекаться от цифровых вычислительных машин, так как они, как я покажу ниже, явятся чрезвычайно важным инструментом при попытке продвинуться дальше в нашем континууме.

Значение исследований в области искусственного разума

В своем введении я высказал уверенность в том, что работы по дальнейшему совершенствованию машин в континууме разумности являются чрезвычайно важными. Сегодняшние вычислительные машины помогают расширять границы человеческих знаний во многих областях, потому что вычислительные машины используются почти во всех научных дисциплинах. (Тот факт, что они используются в исследованиях по созданию машин, наделенных искусственным разумом, означает, что эти исследования питают сами себя.) Применение вычислительных машин во всякого рода исследованиях послужило основным фактором резкого научного подъема, свидетелями которого мы явились в последнем десятилетии. До сих пор главными их показателями являлись скорость и надежность выполнения арифметических операций. Но так как машины движутся вперед в континууме возможностей, то могут появиться новые важные показатели. Например, машина, воспроизводящая или достаточно точно апроксимирующая процесс, протекавший в голове д-ра Уиллиса Вэра и приведший к тому, что Вэр произнес «Фрайзел», но которая работает быстрее и надежнее д-ра Вэра и имеет больший объем памяти,— такая машина может быть спо-

собной увековечить современные темпы развития науки. Без такой машины ученые могут потонуть в результатах исследований, они не смогут усвоить все эти результаты. Даже сегодня многие ученые считают, что часто бывает легче повторить исследование, нежели искать, где оно уже было выполнено и где записаны его результаты. Теперь очевидно, что такая машина («Вэр — Фрайзел») имела бы колоссальное значение для всех технологических отраслей знания.

Огромные затраты на разработку и строительство машин сегодня говорят о том значении, которое мы придаем таким свойствам этих машин, как быстродействие и надежность. А если способности машин будут развиты еще и в других направлениях, то разве не будет это достижением огромного значения? Допустим, что граница, выше которой машина не может подняться, проходит вблизи способностей человеческого мозга на «оси сложности» применяемой методики обработки информации. Так как мы знаем, что человека нетрудно превзойти по быстродействию и надежности и, возможно, по объему памяти, то такая машина приблизилась бы к состоянию «сверхчеловека». Безусловно, это лишь мои размышления, граница может быть в действительности значительно ниже.

Очевидно, значение исследований в области разумных машин не определяется обязательной постановкой цели построить в конечном итоге машину-«сверхчеловека». Если вы признаете, что предел способностей человека конструировать мыслящие машины весьма высок в нашем часто упоминаемом континууме, то, мне кажется, вы должны признать и крайне важное значение исследований, служащих нашему дальнейшему продвижению в этом направлении.

Я работаю над вопросами технического значения исследований в области искусственного разума с точки зрения расширения границ технических и научных знаний. В этом направлении нет особых причин для беспокойства: в США развитие исследований в этом направлении идет бурными темпами. Однако в мировых масштабах в свете нашего технического соревнования с СССР всякое действие в этом направлении оказывается существенным, особенно, по моему мнению, с тех пор, как русские стали отводить исследованиям в области искусственного мозга

большую роль, чем мы. Даже если бы русские не были нашими соперниками на пути к этому «техническому Олимпу», все равно нам стоило бы увеличить темпы.

Подходы к проблеме конструирования разумной машины

Мне кажется, то, что я хочу сказать по этой теме, будет более понятно, если я воспользуюсь в своих объяснениях аналогией. Допустим, нам дали устройство и мы, пронаблюдав за его действиями, убедились, что оно способно разумно себя вести. Нам хотелось бы сконструировать машину, обладающую почти такими же способностями или, еще лучше, бо́льшими способностями. Мы поручаем группе людей изучить основные блоки данного устройства. Эти люди изучают блоки, стремясь понять, как они работают. Люди подают на вход импульсы и наблюдают, что делает каждый блок. Они стремятся понять с точки зрения физики и химии, почему устройство делает именно то, что́ делает. Они также стараются выяснить, как эти блоки функционируют, будучи собраны в узлы (подсистемы).

Тем временем мы собираем вторую группу людей. Она решает данную проблему, рассматривая это устройство как «черный ящик», который они не могут открыть. Эта группа устанавливает, что одни принадлежности данного устройства являются, очевидно, его входом, а другие — выходом. Группа наблюдает это устройство в действии и пытается понять теоретические принципы его работы. Эта группа людей исходит из того, что совсем не обязательно машина, которую им предстоит построить, должна состоять из тех же самых основных компонентов, что и изучаемое устройство. Они считают, что если им удастся понять логику действия существующего устройства, то они смогут воспроизвести эту логику в своей машине, построенной из известных им элементов и блоков.

Эта группа делает предположения относительно логической структуры данного устройства и проверяет эти предположения на имеющейся под рукой вычислительной машине. Эти теоретические догадки сначала бывают очень сырыми, незрелыми и недостаточно хорошо отражают поведение «черного ящика», но с течением времени сходство растет.

Поскольку мы извлекли урок из наших попыток создать летательный аппарат, машущий крыльями, мы учреждаем третью группу людей и поручаем ей изучать «разум» и обработку информации per se и построить научную теорию в этой области. У этих трех групп много общего, и они регулярно обмениваются своими результатами. Они используют в своих исследованиях вычислительные машины. Они объединяют свои усилия для создания лучших вычислительных машин (разумных машин с наибольшим частным интеллектуальности), на которых они проверяют свои предположения. Наконец, эти три группы приходят к общему решению и создают машину, обладающую почти всеми способностями исходной модели. Затем они обращаются к задаче построения машины даже лучше исходной.

В реальной обстановке, в случае изучения человеческого мозга первая группа, изучающая компоненты и узлы, составленные из них, — это не что иное, как группа физиологов. Так же легко найти аналогичные примеры для групп «черный ящик» и «per se». Мне кажется, эта аналогия представляет правдоподобную картину положения вещей в процессе познания человеческого разума.

Теперь несколько слов о положении дел в рассматриваемых областях в СССР.

Русские начали работать над вычислительными машинами позже нас, но уже определенно сократили разрыв. Они придают большое значение развитию вычислительной техники. По-моему, нет оснований считать, что потребность в вычислительных машинах будет служить помехой для работ по созданию разумных машин, хотя машинное время у них сейчас является дефицитом.

В математике русские давно уже заслужили отличную репутацию. В вычислительной математике, я не сомневаюсь, они в общем перегнали Запад. Нашу делегацию и другие еще до нас [25] поразило то обстоятельство, что в области вычислительной техники сейчас работают много выдающихся математиков. К сожалению, многие наши математики смотрят на вычислительную машину как на прославленную логарифмическую линейку, служащую лишь интересам инженеров, или как на дорогой табулирующий прибор, предназначенный для решения канцелярских проблем.

В отношении психологии и физиологии я мало компетентен. Большинство сравнений, которые я слышал или о которых читал, ограничены выяснением первенства. Конечно, эта область n-мерных и одномерных высказываний относительно того, кто из нас впереди, является трудной, если не неразрешимой. Значительно больший интерес представляет вопрос, какую часть айсберга составляет то, что мы видим. Единственная статья, которую я читал и в которой заявляется, что США отстают от Советского Союза,— это статья Грегори Разрана, помещенная в «Сайенс» [26]. Из статьи явствует, что автор весьма обеспокоен состоянием нашей науки из-за реальной возможности прорыва в научной теории об управлении мышлением и действиями человека. Он говорит: «...Необходимость догнать и перегнать СССР в важнейших областях психофизиологии и науки об управлении действиями человека, вероятно, является такой же насущной, как и подобные задачи в физике и теории управления реактивными снарядами».

Я, конечно, не знаю, насколько автор осведомлен о достижениях русских, и не понимаю соотношения между психофизиологией и теорией управления мышлением и действиями человека, но если Разран правильно оценивает соотношение достижений США и СССР в области психофизиологии, то я очень огорчен. Многие из вопросов, которые он затрагивает, совпадают с интересами первой (компонентной) группы из моей аналогии, относящейся к подходам к исследованиям в области разумных машин. Замените слова «управление действиями человека» словами «конструирование разумных машин» — и получится моя концепция.

Я ехал в Советский Союз, будучи уверен в том, что русские уделяют много внимания исследованиям в области искусственного разума. И я все время искал подтверждения этому. Возможно, то, что я искал, оказало влияние на мои выводы относительно того, что я находил. Хочу подчеркнуть, что *меня поразили не какие-либо существенные их результаты, а их явное убеждение в том, что это* важная область исследований.

Слово «кибернетика» прочно вошло в русский язык. Много пишется по этому вопросу в научных журналах и популярной литературе. Отмечается необходимость использовать кибернетику для строительства коммунизма.

Что касается профессиональной литературы, то в 1958 году был учрежден журнал под названием «Проблемы кибернетики». К настоящему времени вышло четыре тома в твердом переплете [27]*. С 1955 года в МГУ учреждаются семинары по кибернетике, которые с 1956 года проводятся еженедельно. Эти семинары имеют целью сближение ученых разных специальностей, объединяемых кибернетикой. Аналогично этому редакторы «Проблем кибернетики» заявляют, что их целью является «унификация научных интересов специалистов, работающих в различных областях знаний, связанных с кибернетикой». По-видимому, существует широкое признание необходимости всенаучного подхода к проблемам кибернетики. Статья за статьей призывает ученых разных профессий объединиться. Какой эффект имеют эти призывы и эти семинары — неизвестно. Но нам сказали, что около 500 физиков обратились к биологическим наукам. Мы разговаривали с И. М. Гельфандом, всемирно известным математиком, ныне работающим в области физиологии. Он начал изучать мозг, но переключился на сердце, которое, он считает, устроено много проще. Со знаниями, полученными при изучении сердца, он вернется к изучению мозга. Нам также рассказывали, что другие математики работают над психологическими и физиологическими проблемами.

При Академии наук СССР существует Научный совет по кибернетике. Советом руководит академик А. И. Берг, и, очевидно, этот Совет подотчетен Президиуму Академии [28]. ...Время от времени можно встретить сообщение об организации новых групп и лабораторий кибернетики.

О сроках

Прежде чем закончить, я хотел бы высказать некоторые соображения по поводу вопроса о сроках. Одно дело — сказать, что совершенствовать дальше способности машин в континууме разумности возможно, другое дело — сказать, когда это будет. Да Винчи и братьев Райт разделяет свыше 400 лет. Но сейчас движение в научном мире про-

* Очевидно, имеется в виду непериодический сборник «Проблемы кибернетики», выходящий под редакцией проф. А. А. Ляпунова. К настоящему времени вышло десять выпусков.— *Прим. ред.*

исходит много быстрее, чем раньше. То, что сейчас достигается за 10 лет, сравнимо с тем, что раньше занимало столетие. Мало кто поверил бы в 1950 году, что не пройдет и десяти лет, как человек запустит на Луну ракету. Недавно по поводу предположения Симона и Ньювела, что машина через какие-нибудь десять лет станет чемпионом по шахматам [29], один из крупных специалистов по информационно-логическим машинам назвал это предсказание консервативным и сказал, что это случится скорее.

Заключение

Я надеюсь, что определение исследований в области искусственного разума как попыток дальнейшего развития машин в континууме разумного поведения уменьшит возможность семантических недоразумений при обсуждении таких исследований. Я уверен, что эти исследования имеют большое значение для нашей страны, и мы должны умножить наши усилия в этом направлении. Для этого необходимо привлечь больше исследователей соответствующих профессий. Успех наших усилий будет зависеть от того, насколько нам удастся сблизить представителей различных специальностей и как много квалифицированных специалистов удастся привлечь к этим исследованиям.

ЛИТЕРАТУРА

1. R. A MacGowan, Letter to the Editor, «Science», July 22, 1960.
2. S. Butler, Erewhon and Erewhon Revisited, Modern Library No. 136, Rondom House, New York, 1933.
3. B. V. Bowden, (Ed.), Faster Than Thought, Pitman, London, 1953.
4. N. Wiener, Some Moral and Technical Consequences of Automation, «Science», May 6, 1960.
5. M. Taube, Letter to the Editor, «Science», August 26, 1960.
6. R. V. Andree, Programming the IBM 650 Magnetic Drum Computer and Data Processing Machine, Holt, 1958.
7. E. M. Hugh-Jones, Automation Today, «Automation in Theory and Practice», Oxford, 1956.
8. J. Kister, P. Stein, S. Ulam, W. Welden and M. Wells, Experiments in Chess, «J. of the Association for Computing Machinery», April 1957.

9. A. N e w e l l, J. C. S h a w and H. A. S i m o n, Chess-Playing Programs and the Problem of Complexity, «IBH J. of Research and Development», October 1958.
10. A. N e w e l l, J. C. S h a w and H. A. S i m o n, Empirical Explorations of the Logic Theory Machine, «Proc. of the Western Joint Computer Conference», February 1957.
11. J. H. T r o l l, The Thinking of Men and Machines, «Atlantic Monthly» July 1954.
12. K. R. S h o u l d e r s, On Microelectronic Components, Interconnections, and System Fabrication, «Proc. of the Western Joint Computer Conference», May 1960.
13. A. M. T u r i n g, Computing Machinery and Intelligence, «Mind», October 1950; «The World of Mathematics», Vol. 4, p. 2099.
14. H. L. G e l e r n t e r, Realization of a Geometry Theorem Proving Machine, «Proc. of the International Conference on Information Processing», UNESCO, Paris, France, 1960.
15. A. L. S a m u e l, Some Studies in Machine Learning Using The Game of Checkers, «IBM J. of Research and Development», 3, No. 3, July 1959.
16. F. T o n g e, A Heuristic Program For Assembly Line Balancing, The RAND Corporation paper P-1993, May 18, 1960.
17. L. A. H i l l e r and L. M. I s a a c s o n, Experimental Music, McGraw-Hill, 1959.
18. G. L. G o o d w i n, Digital Computers Tap Out Designs for Large Motors, Fast, «Power», April 1958.
19. B. G o l d, Machine Recognition of Hand Sent Morse Code, «IRE Trans. of Information Theory, March 1959.
20. A. N e w e l l, J. C. S h a w and H. A. S i m o n, Problem Solving in Humans and Computers, The RAND Corporation paper P-987, December 7, 1956.
21. M. M i l l i g a n, Machines Are Smarter Than I Am!, «Data Processing Digest», October 1959.
22. J. M c C a r t h y, Programs With Common Sense, «Proc. of the Symposium on the Mechanization of Thought Processes», NPL, November 1958; Her Majesty's Stationery Office, 1959.
23. M. M i n s k y, Methods of Artificial Intelligence and Heuristic Programming; «Proc. of the Symposium on the Mechanization of Thought Processes», NPL, November 1958; Her Majesty's Stationery Office, 1959.
24. M. V. W i l k e s, Automatic Digital Computers, Wiley, 1956.
25. J. W. C a r r, et al, A Visit to Computation Centers in The Soviet Union, «Communications of The Association for Computing Machinery», June 1959.
26. G. R a z r a n, Soviet Psychology and Psychophysiology, «Science», November 14, 1958.
27. Сборники «Проблемы кибернетики» под ред. А. А. Ляпунова, Физматгиз.
28. А. И. Б е р г, Кибернетика и общество, «Экономическая газета», 12 июня 1960.
29. A. N e w e l l and H. A. S i m o n, Heuristic Problem Solving: The Next Advance in Operations Research, «J. of the Operations Research Society of America», 6, No. 1, January-February 1958.

Содержание

М. ТАУБЕ

Вычислительные машины
и здравый смысл

Редактор *А. Лукина*

Художник *Г. Дауман*

Художественный редактор *Л. Шканов*

Технический редактор *М. Сафронович*

Сдано в производство 4/VI 1964 г.
Подписано к печати 6/X 1964 г.
Бумага 84×108/32=2,9 бум. л.
9,4 печ. л., в т/ч. вкл.
Уч.-изд. л. 9,9.
Цена 60 к. Зак. 240

ИЗДАТЕЛЬСТВО «ПРОГРЕСС»

Москва, Зубовский бульвар, 21

Московская типография № 16
«Главполиграфпрома» Государственного
комитета Совета Министров СССР
по печати. Москва, Трехпрудный пер., д. 9


Цена 60 к.